Title
Cultural complexity and complexity evolution

Journal
Adaptive Behavior, 28(5)

ISSN
1059-7123

Authors
Read, D
Andersson, C

Publication Date
2020-10-01

DOI
10.1177/1059712318822298

Peer reviewed

Cultural Complexity and Complexity Evolution

(accepted for publication in *Adaptive Behavior* on December 10, 2018)

Dwight Read
Department of Anthropology
UCLA
Los Angeles, CA 90035 USA

Claes Andersson
Department of Space, Earth and Environment
Chalmers University of Technology
412 96 Göteborg Sweden.

Abstract
We review issues stemming from current models regarding the drivers of cultural complexity and cultural evolution. We disagree with the implication of the treadmill model, based on Dual-Inheritance Theory, that population size is the driver of cultural complexity. The treadmill model reduces the evolution of artifact complexity, measured by the number of parts, to the statistical fact that individuals with high skills are more likely to be found in a larger population than in a smaller population. However, for the treadmill model to operate as claimed, implausibly high skill levels must be assumed. Contrary to the treadmill model, the risk hypothesis for the complexity of artifacts relates the number of parts to increased functional efficiency of implements. Empirically, all data on hunter-gatherer artifact complexity support the risk hypothesis and reject the treadmill model. Still, there are conditions under which increased technological complexity relates to increased population size, but the dependency does not occur in the manner expressed in the treadmill model. Instead, it relates to population size when the support system for the technology requires a large population size. If anything, anthropology and ecology suggest that cultural complexity generates high population density rather than the other way around.

Introduction
In this article we review some of the issues stemming from current models relating to the drivers of cultural complexity and cultural evolution. The issues we raise are not with the form of the models, but with background assumptions that have been made regarding what constitutes cultural complexity and cultural evolution. Additional issues arise through the empirical evidence advanced to support theoretical models regarding the evolution of cultural complexity, especially in small scale, hunter-gatherer societies.

In particular, we take issue with a family of models and empirical investigations deriving from the so-called "treadmill model" – the name Kline and Boyd (2010) have given to the mathematical model developed by Henrich (2004). This model has played a prominent role in this literature with its provocative implication that the interaction population size – the census population of those who reside together plus those with whom they interact on a regular basis, or what Henrich (2004) refers to as the effective cultural population size – is the prime driver of cultural evolution leading to increased cultural complexity (Henrich, 2004; Powell, Shennan & Thomas, 2009; Shennan 2001). The empirical evidence said to support this model and its derivatives have been archaeological (e.g., Powell, Shennan & Thomas, 2009), ethnographic (e.g., Henrich 2004; Kline & Boyd 2010), and experimental (e.g., Derex et al., 2013). Archaeological data provide the time depth required for tracing out the evolutionary pattern of time-based changes in the properties and features of human societies, but tend to be limited to the material remains that preserve through time. Ethnographic observations made on living groups provide a broader spectrum of data, thus allowing for a more complete and more nuanced account of the cultural repertoire of a group, but, for the most part, lack the time depth needed for evolutionary observations. Experimental data provide the means for examining, under controlled conditions, the interaction between complexity and group size – the primary components of the "treadmill model" – by identifying factors that affect the complexity of what a small group can produce.

The issue that arises with empirical evidence advanced to support the treadmill model is largely one of interpretation: Do observed properties, and changes in those properties through time, match the assumptions of theoretical models so that concordance between observed change and model prediction constitute support of the theory underlying the model? What we will show in this paper is that the background assumptions required for the theoretical models that claim to establish the interaction population size as a primary driver of change in cultural complexity, hence an important driver for cultural evolution, are not supported by the empirical data brought forward and said to support the theoretical models. A critical background assumption shown to be invalid by this discordance is the assumption of the treadmill model that increase in cultural complexity is driven by increase in the expected number of highly skilled individuals in a population as the population size gets larger. Yet already with a population size of n = 500, the expected value for the number of individuals in the 98th percentile for normally distributed skill levels is 10, thus a hunter-gatherer society of size n = 500 will already be likely to have several individuals with a high skill level. The limiting factor for increase in cultural complexity in hunter-gatherer groups, then, is not, as claimed in the treadmill model, the likelihood of the absence of skilled individuals in a hunter-gatherer group due to a small population size, but whether the payoff in investing in a culturally more complex system justifies the cost of so doing.

Two other reasons why the relationship between empirical data and the treadmill model is problematic are as follows. Firstly, the treadmill model for increase in cultural complexity is based on an overly simplistic view of what actually constitutes complexity vis-à-vis human culture. What do the empirical tests aimed at validating the treadmill model really measure and so with what is a test for correlations with cultural complexity being made? Secondly, even granted that the empirical tests are legitimate, the complexity of implements made by hunter-gatherer groups does not correlate with the interaction population size, as claimed in the treadmill model, but with mobility and food procurement risk (see Read, 2008). What is actually being measured is the complexity of individual tools, which reflects investment in their production, curation, transportation and use effectiveness as part of coping with food procurement risk. This investment is subject to a trade-off between risk and mobility since material culture tends to "entangle" those that rely upon it (Hodder, 2012). This tradeoff leads to a subdivision of the cultural evolutionary design space of cultural societies considered as wholes (Taylor, 2010): Do groups invest in the ability to move to other places where a simpler technology works (what Binford [1980] refers to as a forager strategy), or do groups invest in a more complex technology that permits more intensive usage of the area where the group is located (what Binford [1980] refers to as a collector strategy)? This difference in investment and the predicted difference in the complexity of technology is precisely what characterizes the way tool complexity relates to risk and mobility for hunter-gatherer societies (see Figure 13, below). In other words, what is primarily measured through tool complexity is an indicator of how the complexity of culture is distributed *within* cultural systems. While that is an important question in and of itself, tool complexity cannot be used as a direct proxy for overall cultural complexity, as asserted by the proponents of the treadmill model.

Expansion of Models of Evolution to Include Cultural Traits

Over the past several decades, mathematical models of biological evolution have been expanded to include not only biological traits but also so-called cultural traits. In the latter half of the 20th century, researchers such as Campbell (1960, 1964, 1974), Cavalli-Sforza and Feldman (1981), and Boyd and Richerson (1985) realized that, with humans and human societies, trait transmission involves phenotypic as well as genotypic traits when the phenotypic trait is not simply the developmental expression of an individual's genotype but is, for example, behavior arising through the neurological processes of the brain. Boyd and Richerson refer to phenotypic trait transmission like this as cultural trait transmission – which unfortunately conceals an important distinction between cultural systems and what they are made of. Cultural transmission occurs through enculturation, the process by which cultural systems, as a whole, become part of the cultural identity that a newborn takes on as he or she develops through being immersed and raised in an ongoing cultural community, much like a newborn takes on a language through being immersed and raised in an ongoing community of language speakers. What they refer to as a cultural trait is more commonly – and specifically – referred to as a tradition, which may involve beliefs, objects or customs that are transmitted through time by being taught by the members of one generation to the members of the next generation (Shils, 1981). Hence, rather than using their reference to cultural trait transmission, we will refer instead to a traditional transmission mode.

A traditional transmission mode for behavior learned in a social context by observing the actions of others is today known to be widespread (see, e.g., Dean et al., 2014; Galef, 1992; Galef & Laland 2005; Laland & Janik, 2006). *Homo*, however, came to develop an enigmatic and extensive cumulative cultural system based on a much-buttressed version of this traditional transmission of behavior along with the ideas underlying these behaviors – including a whole suite of derived cognitive and psychological adaptations for learning and, not least, teaching (see, e.g., Barrett & Henzi, 2005; Burdett et al., 2017; Castro & Toro, 2014; Csibra & Gergely, 2011; Gärdenfors & Högberg, 2017; Kline, 2015).

The difference in trait transmission between genetic traits and traditional traits has been implemented in Dual-Inheritance Theory (DIT) (Boyd, 2018; Boyd & Richerson, 1985). DIT was introduced as a means to consider the interplay between traditional and genetic transmission. The unifying idea of DIT is that the difference between the evolution of cultural and of biological features lies mainly in differences between these two types of transmission, not in the nature of the features.

Most would agree that reproductive fitness is a more complex matter for a traditionally transmitted trait than for a genetically transmitted trait. Two aspects in which this difference is displayed include the mode of transmission and the generation of variation.

The genetic part of DIT is focused primarily on differences arising through vertical transmission. The view that genetic transmission is inherently and essentially simple and vertical (see, e.g., Sterelny, 2011), though, is now considerably more nuanced due to dramatically improved knowledge of microbial evolution, and is now seen to be more like the transmission of traditions than was previously realized. It is still true that the several alternative pathways of transmission that exist for traditional transmission are qualitatively different from the transmission pathways that are relevant to biological transmission. Transmission of traditions comes with potentially different criteria for "passage," including not only vertical (from parent to offspring) transmission, but transmission that is horizontal (between individuals of the same generation), and oblique (between individuals of different generations). Other criteria affecting the fitness of tradition transmission, but not genetic transmission, include social relationships that exist among individuals in a group, such as gregariousness, conformism, and, in humans, a tendency of individuals to use prestigious individuals as role models (Boyd 2018; Henrich & Gil-White, 2001).

Trait variation involves *development* in the broad sense of how hereditary information is transformed into functional phenotypic organization with regard to the environment that contributes to differential fitness. This has been a neglected factor in evolution in the Modern Synthesis where the foundations of modern evolutionary theory that DIT departs from took shape (e.g., Laubichler & Maienschein, 2007, 2013; Laubichler & Renn, 2015), and is due, to a large extent, to a lack of empirical knowledge about these processes until recently. In the context of cultural evolution, development has to do with culture as an integrated system, as Kroeber (1919) observed with regard to fashion trends: "The reintroduction of the train in 1863, the invention of the Grecian bend in 1872, may now be looked upon as the product of the dress styles that preceded them or of other cultural factors affecting style, more justifiably than they can be attributed to the talent of a specially gifted mind and hand" (p. 260) (see also Andersson, Törnberg & Törnberg, 2014a; D'Errico & Banks, 2013). Culture as an integrated system is an aspect that DIT – which is a fundamentally microevolutionary theory – tends to disregard. For DIT, evolution is where "one choses from a pool of variants … and the individual-level processes of selection determine the success, at the population-level, of the variants" (Acerbi & Mesoudi, 2015, p. 483). The disregard of the integration of cultural systems is reflected in how cultural complexity is perceived in DIT-related research – which contrasts both with modern biological debates on the same topic (e.g., Marcot & McShea, 2007; McShea, 1991, 2000) and with the anthropological view of culture complexity. (Discussions about macroscopic cultural complexity in the context of cultural evolution – a separate topic in its own right – include Andersson 2013, p. 90; Andersson, Törnberg & Törnberg, forthcoming; Querbes, Vaesen & Houkes , 2014; Read, Lane & van der Leeuw, 2009).

Definitions of Culture and Cultural Traits

Defining cultural traits through phenotype transmission successfully expanded the scope of evolutionary models beyond the evolutionary consequences of genotype transmission. This has made it evident that phenotypic trait transmission has evolutionary consequences that fall outside of the scope of the evolutionary consequences of genotypic trait transmission. The extension of evolutionary theory to also include trait transmission has, however, not been achieved without a cost. The extension of the concept of a trait to include not only genetically but also traditionally transmitted traits has reduced culture to simply be that part of an individual's phenotype that is transmitted traditionally from one individual to another. In this perspective, cultural traits are distinguished from biological traits only by the mode of transmission and not by what constitutes the domain of culture.

This runs counter to how anthropology has normally viewed culture, which is as socially implemented systems of ideas, world views, beliefs, and the like shared by the members of a community or society (Firth, 1951). The definition of culture given by Edward Burnett Tylor in 1929[1871], a definition that is still fundamental to how anthropologists view culture (Avruch, 1998; Ping, 1999), remains central to ethnographic research and writing (Herbert, 1991):

> Culture or Civilization, taken in its wide ethnographic sense, is that *complex whole* which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits *acquired by man as a member of society* (p. 245, emphasis added).

For Tylor, culture is composed of two essential parts: first, what culture is – a "complex whole"– and second, how it is obtained – "acquired by man as a member of society."

It is, however, only the second part of Tylor's definition, how culture is obtained, that was taken up by Boyd and Richerson (1985) as the defining characteristic of what distinguished a cultural trait from a genetic trait. This puts what constitutes culture in sharp contrast with the first part of Tylor's definition, for culture as a "complex whole" does not relate to the mode of transmission, but to integrating together its constituent parts: "a culture is more than a fortuitous assemblage of traits; each culture possesses, in addition to its trait content, a unique organization in terms of which its distinct components are significantly related to one another" (Hoijer, 1948: 338). This implies that the nature of "religious systems, social functions, structures of kinship, and modes of production meant that the analysis needed to begin at the level of society or culture as a whole" (Bubandt & Otto, 2010: 8), hence with it comes a focus on what is known in common and is shared by community members (Ping 1999).

The word "complex" in the phrase, *complex whole*, signals that culture is *not* a single, indivisible entity in the sense that the term *holism*, in anthropological theory, came to be "geared toward asserting bounded, static, homogeneous wholes … [and] prone to be used in totalizing ways" (Bubandt & Otto, 2010, p. 2, 9). Thus, the phrase, complex whole, implies that culture is more than the sum of its individual parts (Wallis, 1930) by being an integrated whole (Ferraro, 1998) in which changes in one aspect leads to changes in all other aspects (Hoijer, 1948). Consequently, evolution of culture has to do with matters such as change in the integration of culture, changes in the idea systems comprising culture that are known to, and shared by, community members (Leaf & Read, 2012), and changes in the form of cultural organization (Lane et al., 2009). Cultural evolution is not simply change in the frequency of a cultural trait measured over a population of individuals.

The word "complex" also suggests that culture is neither a single whole in the sense of a single system with numerous parts, each well designed and carefully linked to each other in the system, yet no one of which bears a reflection of the whole for which it is a part, much like the parts of clock do not individually express the whole, for the clock is seen only in the assembly of the individual parts. Nor is culture a whole that simply emerges from the interaction of its individual parts (*contra* Smaldino, 2014; see also Read, 2014b). Were this the case, change in one part need not have any effect on other parts or on what emerges, but for culture, as the linguist Harry Hoijer (1948) points out: "[c]hanges in one aspect of a culture must inevitably result, sooner or later, in changes in all other aspects" (p. 338). Thus, culture, as a "complex whole," is *more* than a property emergent from the interaction of simpler parts and *less* than a rigid system of interconnected parts. It is somewhere in between (Andersson, Törnberg & Törnberg, 2014b, forthcoming) – a quality that is reflected for example in language in that the communicative aspect of language is expressed through speech formed from language as a complex whole through its grammar, yet the communicative aspect of language is also conveyed by individual words.

A classic example of a DIT approach to cultural phenomena is provided by the argument made by Durham (1991) for the universal occurrence of incest taboos forbidding sexual relations between parent and child or between siblings. According to Durham's argument, the genetic component of the incest taboos consists of the biological consequences of inbreeding and the cultural component is the taboo itself. For the non-human primates, inbreeding avoidance can arise through selection for sex-biased philopatry and, at least for chimpanzees, by biological sons avoiding sexual copulation with their biological mothers. Presumably, with the increase in the age of sexual maturity during hominin evolution leading to *Homo sapiens*, inbreeding avoidance alone was not sufficient for inbreeding avoidance (Turner and Maryanski, 2009). From a cultural perspective, Durham argues that the evolutionary transition from hominins to *Homo sapiens* includes the formation of the brain functions and language abilities that made possible the development of culture and its social transmission in human societies (p. 155). Through their observations, he argues, our *Homo sapiens* ancestors would have seen the deleterious consequences (but not the mechanism) of inbreeding depression. Some groups would hit upon the idea of avoiding the matings associated with deleterious effects, and some of these groups, he argues, made these matings taboo. The taboos would have been passed on through social transmission, leading to cultural group selection favoring these groups, thereby giving rise to the universal incest taboos that, according to Durham, are neither the consequence of a genetic component alone nor a cultural component alone.

It should be noted that it is not Durham's DIT argument, but a non-DIT genetic argument that has been popularized as the supposed explanation for the incest taboos. The alternative argument relies on Westermarck's (1922) claim that individuals raised together (regardless of genetic affinity) will abhor having sex with each other. The incest taboo is then postulated to be the cultural expression of this abhorrence.

This argument has been heavily criticized on several grounds. Firstly, El Guindi (2015) notes that Westermarck references Mundt (1857) to the effect that "it was not uncommon for French fathers to live in concubinage with daughters" (1922, p. 200) and that "the French nature is not repelled to the same degree as the German by the idea of sexual unions between persons nearly related by blood" (1922, p. 200), yet Westermarck does not consider these observations to contradict his claim about sexual abhorrence since he considers them to be "quite exceptional" 1922, (p. 201). Secondly, all of the data that supposedly support Westermarck's argument (*kibbutzim* data [Sepher, 1983], Lebanese cousin marriage data [McCabe, 1983], and Taiwan minor marriage data [Wolf, 1995]) have been shown to have more plausible, alternative interpretations (see El Guindi & Read, 2012; Leavitt, 2005, 2012; Read, 2014a; Shor & Simchai, 2009, 2012, among others) or, in the case of brother-sister marriages in Roman Egypt (Middleton, 1962; Scheidel, 2005), the data directly contradict Westermarck's claim. Thirdly, Arthur Wolf, one of the main proponents for the Westermarck hypothesis as the basis for incest taboos, now argues that "incest avoidance and the incest taboos [are] different things with different origins" (Wolf, 2014b; see also Wolf, 2014a).

Durham's DIT argument is also problematic, but for two different reasons. Firstly, on a more technical note, he (and others) assume incorrectly that inbreeding depression is a feature of mating with close biological relatives rather than the consequence of a shift from outbreeding to inbreeding (Shields 1982). Regardless of the mating pattern, at genetic equilibrium the rate of occurrence of deleterious traits due to recessive alleles is a function of the mutation rate for deleterious alleles and is independent of the mating pattern (Read 2014a). Secondly, he assumes the incest taboo can be considered in isolation from the larger system of cultural concepts for which it is a part. How the incest taboo is part of a more encompassing cultural system is made evident in Lévi-Strauss's (1949) observation that:

> "[t]he prohibition of incest is where nature transcends itself. It marks the formation of a new and more complex type of structure and is superimposed upon the simple structures of physical life through integration, just as these themselves are superimposed upon the simpler structures of animal life. It brings about and is in itself the advent of a new order" (p. 25).

Lévi-Strauss's argument may be fleshed out by noting that the kinship relations making up the kinship systems that play a central role in all human societies are derived from an atomic structure formed from the procreation based relations of parent-to-child and sibling-to-sibling and the culturally determined relation of spouse-to-spouse (Chit Hlaing & Read, 2014) and referred to as a Family Space by Read and co-workers (see Read, Fischer & Chit Hlaing, 2011). Critically, the coherency and logical consistency of the Family Space of primary relations – hence the coherency and logical consistency of the kinship relations derived from this structure – depends on making parent-child and sibling-sibling marriages "unthinkable," a status achieved culturally by making taboo sexual relations between parent and child and between sibling and sibling (Read, 2018).

Another crucial aspect of what is meant by culture is expressed in Ward Goodenough's widely referenced observation that "a society's culture consists of whatever it is one has to know or believe in order to operate in a manner acceptable to its members, and to do so in any role that they accept for any one of themselves" (Goodenough, 1964, p. 36). Goodenough also brings out, in a different way, the key point that culture is not the sum of individual traits since a key aspect of culture is the way it makes possible the coordinated interplay of individuals within the social group for which they are members. The coordinated interplay depends on mutual understanding by group members of what is considered to be proper, or acceptable, behavior. This means that, from the perspective of an individual, the functionality of culture does not arise through a trait expressed at the individual level but through an individual's interaction with the social group for which the individual is a member: "Individuals acquire knowledge about … culture by interacting with one another; … these meanings may be modified as they interact with others" (Greenberg, 1961, p. 10).

A critical example is a kinship terminology composed of kin terms generated from the relations making up the Family Space. For English speakers, these kin terms are father, mother, brother, sister, son, daughter, uncle, aunt, husband, wife, … and so on. For other groups, the kin terms differ not only due to a language difference, but also by virtue of kinship ideas expressing which kin relations are applied to which persons. English speakers, for example, refer to a man as uncle when that man is the father's brother, the mother's brother, the father's sister's husband or the mother's sister's husband of the speaker. In many terminologies, though, different kin terms are used for these genealogical criteria according to whether the individual in question is related to the speaker through speaker's mother or through speaker's father; that is, maternal relatives are identified by different kin terms than paternal relatives, whereas the maternal/paternal distinction is not part of the English (and most European) kinship terminology.

Just as each group member is knowledgeable of the language that the group uses for communication, each group member is knowledgeable of the kinship terminology that is part of the cultural repertoire of one's group and expresses the kinship relations group members have to one another. If, in line with DIT, we were to consider each kin term to be a cultural trait of an individual, the problem immediately arises that there is no functionality that occurs for an individual from individually knowing a kin term. For example, suppose an English speaker decides to use the expression *unclemo* for mother's sister's husband and *unclefa* for father's sister's husband. If only the speaker uses these new terms, they would be meaningless expressions for anyone else, hence there is no functionality accruing to this person through having the equivalent of a mutation in a trait at the individual level. Of course, neologisms can be introduced and then spread in the population, but the manner in which they do so is constrained by the organization of the language and, semantically, how it is used. The point is that the fitness of a term is not inherent to the term itself. The criteria constraining how the existing structure of kinship terms may change derives from the way kinship terminologies are logically structured as a system with a generative logic for the structure and organization of the kin terms making up the kinship terminology (discussed in Leaf & Read, 2012; Read, 1984, 2001, 2007b; Read, Fischer & Chit Hlaing, 2014, among other references). Thus, a speaker's attempt to use the terms *unclemo* and *unclefa* would be considered incorrect by other users of the English Terminology since these terms are not consistent with the generative logic of the terminology. Of course, just as speakers of a language may know that a syntactically incorrect phrase does not sound correct but cannot articulate why the phrase is syntactically incorrect, culture-bearers may know that the mutated kin term is not valid without being able to express the underlying logic that makes it an invalid term. For example, for English speakers the construction "cousin-in-law" is syntactically valid, and has for them the meaning "spouse of speaker's cousin," (or possibly "cousin of speaker's spouse"), but the English kinship terminology does not have a kin term whose meaning is "spouse of my cousin" (Hage, 1997). A Google search on "cousin-in-law" shows the ambiguity of this candidate for a kin term. Some English speakers consider cousin-in-law to be a kin term and others do not.

The underlying logic for why "cousin-in-law" is not a kin term for English speakers derives from kinship ideas that are incorporated in the English kinship terminology. One of these is the universal idea of reciprocity of kin terms: If one person has a kin term relation to another person then the latter person has a kin term relation to the first person. If we consider the relationship of a kinship terminology to its constituent kin terms to that of a whole to its parts, the reciprocity property neither occurs at the level of the part nor does it emerge from properties at the level of the part. Rather, it is a concept at the level of the whole that then induces a property at the level of the part. From the reciprocal property it follows that if cousin-in-law is the kin term denoting the kinship relation, spouse of cousin, then the reciprocal of the kin term cousin-in-law must be a kin term in the English kinship terminology as well. Since spouse of cousin = spouse of grandchild of grandparent, the reciprocal of cousin-in-law would be grandchild of grandparent of spouse, but grandparent of spouse is not a kin term for English culture-bearers, hence the reciprocal of cousin-in-law would not be a kin term and so cousin-in-law is not recognized as a kin term since to do so would violate the reciprocity principle. Just as languages have a grammar – a property of the whole – that determines at the level of the part (the words) what are syntactically correct and incorrect utterances, the generative logic underlying a kinship terminology determines what are valid kin terms for that kinship terminology.

At first glance the constraints imposed by a grammar for syntactically correct sentences, or the generative logic of a kinship terminology for what are logically valid kin terms, appear to be comparable to the DIT notion of guided variation. Guided variation refers to the situation where a trait being learned by an individual is modified and the modified trait is then transmitted phenotypically to other individuals (Acerbi & Mesoudi, 2015; Boyd and Richerson, 1985). The modification of the trait might be interpreted as a constraint imposed by the learning process, thereby causing some possible outcomes of the learning process to be more likely than others. Viewed as imposing a constraint, guided variation could be said to be analogous to a grammar or a generative logic with an imposed constraint in the form of syntactic correctness or generative validity, respectively. However, there is a fundamental difference between guided variance and grammars or generative logics. With guided variation there is no frame of reference for evaluating whether the trait modified through learning is "correct" in the sense of whether it fits into, or is consistent with, an integrated system of parts. In addition, a grammar or a generative logic is not the source for a modified phrase or kin term and the selection imposed by the constraint of a grammar or a generative logic relates to whether the phrase or kin term will be accepted as a syntactically valid construction, not whether the phrase or kin term will increase in frequency through phenotypic transmission.

This difference can be seen with the example of the expression, cousin-in-law. Guided variation would refer, firstly, to the person learning the English kinship terminology who then introduces the expression, cousin-in-law, under the assumption that the -in-law suffix is used, in general, to denote a kin relation by marriage, and secondly, to the spread of the expression, cousin-in-law, through phenotypic transmission. From the perspective of the generative logic of the English kinship terminology, however, the expression, cousin-in-law, is rejected as a kin term since it violates the (universal) reciprocity principle for kin terms. Thus, guided variation and the generative logic of the English kinship terminology lead to non-comparable outcomes and so they are not analogous processes.

It is here that Tylor's notion of a whole comes into play. The terminology is not simply a collection of kin terms, each selected in some manner with regard to an external fitness function (whether genotypic or phenotypic in form) but is a conceptual structure with internal rules regarding both syntactic organization and the generation of kin terms from primary terms expressing the relations of the Family Space. How the relations making up the Family Space are instantiated by a particular group gives rise to the various forms that a family may take on (e.g., two parent heterosexual families, two parent homosexual families, single parent families, polygynous families, polyandrous families, extended families, matrifocal families, and so on), depending on the particular cultural context regarding what constitutes a family.

Changes do occur in cultural idea systems like kinship terminologies, but directionality in changes in cultural idea systems are not determined by the fitness benefit accruing to the individual with a changed trait, but by changes in the functionality of the idea system as a whole in providing, in the case of kinship terminologies, the conceptual basis upon which members of a group are conceptually understood by their members to form a social group and how this relates to the functioning of the group as a social unit. Changes in cultural idea systems also have developmental consequences through cascade effects entailed by organizational change (Andersson, Törnberg & Törnberg, 2014a; Lane, 2016; Wimsatt, 2013). For example, American cultural ideas about marriage have changed from an earlier notion that marriage primarily has to do with family formation and establishing the responsibility of a man (the groom) for the well-being of a female and her children, to the present notion that marriage provides public acknowledgment of the love one person has for another. This change has had the cascade effect of legitimizing same sex marriages in American law once it was also accepted that the love emotion, now seen as the basis for marriage, can occur between individuals of the same sex (Read, 2017).

For kinship terminologies, rather than functionality being derived from kin terms viewed as individually expressed cultural traits, functionality derives from being part of an ensemble of individuals jointly knowledgeable about kin terms expressing a system of relations that define for them the kin relations they have to one another, along with rights and obligations associated with those kinship relations. These rights and obligations are expressed, according to Meyer Fortes (1969), through *prescriptive altruism*:

> [K]inship … is associated with rules of conduct whose efficacy comes, in the last resort, from a general principle of kinship morality that is rooted in the familial domain and is assumed everywhere to be axiomatically binding. This is the principle of *prescriptive altruism* which I have referred to as the principle of kinship amity…. Kinship predicates the axiom of amity … [and] kinsfolk are *expected* to be loving, just and generous to one another and not to demand strictly equivalent returns from one another" (pp. 231-232, 237, emphasis added).

In Goodenough's terms, an individual must be sufficiently knowledgeable about the kinship terminology and expected behavior on the part of kinsmen in order for that person to be accepted as a kinsman by the group members, thereby benefitting from the functionality that derives from being part of a group of persons who are kin to another and who act in accord with their cultural knowledge regarding the behavior expected of kinsmen, such as Fortes's notion of prescriptive altruism. However, contrary to the notion of culture expressed in Tylor's definition and its dependency on a social context for its functionality to be realized in the way Goodenough indicates, the DIT view of the relationship between biology and culture considers the domain of biological traits and of cultural traits to differ by the means of transmission and not by the nature of what constitutes the cultural domain in comparison with the biological domain.

The DIT use of the mode of transmission to define cultural traits implies that artifacts are considered to be cultural objects in a very particular way. It suffices that the artifact reoccurs across generations as the instantiation of ideas that are transmitted as traditions. Hence the phenotype – e.g., the instantiation of an idea regarding the shape and form for an artifact – is transmitted from one artisan to another artisan when an artifact is made in accordance with this idea. The artifact made by one artisan can serve as a stimulus for producing the same shape and form of an artifact by a second artisan. Viewing artifact production by artisans in this manner allows the artifact to be considered as the material and observable representation of unseen ideas about the form and shape an artifact should have. In this framework, cultural evolution is then measured by, for example, change in the form and shape of artifacts and cultural complexity is equated with the complexity of the artifact. A major difficulty with this scenario, though, is that the artifact is not, itself, a cultural trait.

Artifacts and the Definition of Culture

In the first half of the 20th century, archaeologists worked out the relationship between artifacts – the material remains from past societies and recovered by archaeologists through excavation of past habitation areas – and culture (see Read, 2007a). Archaeologist Irving Rouse set forth what became foundational ideas for American archaeologists regarding the connection between artifacts and culture. Rouse set out a series of propositions regarding artifacts, the first of which is especially pertinent here. In his first proposition, he made it clear that the ensemble of artifacts produced by group members do not constitute the culture of a group: "1. Culture does not consist of artifacts. The latter are merely the results of culturally conditioned behavior performed by the artisan." (1939, p. 15). Rouse based his argument on Tylor's definition of culture as a "complex whole." According to Rouse, what constitutes culture, though, is not the physical object, but the concepts and ideas underlying the production of an artifact. In a similar vein, the eminent archaeologist Walter Taylor comments that material culture "consists only of objectifications of culture and does not constitute culture itself" (1948, p. 100). Likewise, the eminent British anthropologist Edmund Leach opines: "I shall here use 'culture' in Tylor's narrower sense, while distinguishing the material part of the cultural heritage as 'the products of culture'" (1965, p. 24).

Rouse and others with similar viewpoints were not rejecting the commonly held notion of artifacts being part of "material culture," but were considering that the latter was too limited a notion of how artifacts relate to culture. Artifacts are not simply another kind of culture – material culture – but are the instantiation of an idea system held by the artisans and users of artifacts, and the goal of the archaeologist is to abduce that idea system from the properties of the artifacts made in accordance with the concepts and ideas shared by artisans as part of their culture. Rouse recognized that a primary distinction needs to be made between those features of an artifact that are a consequence of culture; that is, are the consequence of ideas held in common by artisans who are part of the same culture regarding ideas about the characteristics of artifacts, and those features of artifacts that are specific to the artisan who made the artifact. Rouse also recognized that some features are idiosyncratic and are neither the consequence of culture nor the consequence of ideas held by individual artisans. How these ideas can be implemented systematically and objectively is discussed in Read (2007a). And, although Tylor's concept of culture has been critiqued and revised by both archaeologists and anthropologists, his "old-time culture concept still plays an integrating role as a central reference point even for the radically revisionist anthropologists, for whom it is variously a bête noire, a punching bag, or a springboard to alternative perspectives on the human condition, past and present" (Watson, 1995, p.690).

Although constraint, development and sources of variation are debated in the context of DIT (e.g., Eerkens & Lipo 2005, 2007; Lycett & von Cramon-Taubadel, 2015; O'Brien et al., 2010; O'Brien & Bentley, 2011), the theoretical substrate is still fundamentally derived from a model of something else, namely biological trait transmission. So, the focus, in practice, reflects old and partially obsolete biological considerations. For example, any consideration of variational constraints will focus on features that are intrinsic to the artifact itself (such as properties of the raw material) or to the processes by which it is transmitted.

Definition of System Complexity

Like pornography, we know complexity when we see it, but it is hard to define precisely what we mean by complexity of a system. One useful distinction is that between dynamical and organizational complexity (e.g., Andersson, Törnberg & Törnberg, 2014; Andersson & Törnberg, 2018; Érdi, 2008), where the former corresponds to the massively parallel interaction between large numbers of entities (such as flocks of birds, traffic, biological populations, and so on), and the latter corresponds to what we might term "complicatedness." That is, the type of organization prevalent within adapted functional wholes (such as organisms and machines) – corresponding to Simon's concept of near-decomposability (Simon, 1962; Wimsatt, 1975) with a level-hierarchical organization of modules that combine high internal integration with high external separation. The former types of systems may, or may not, have functions (e.g. collective ant behavior *contra* galaxy formation), while the latter are usually adapted systems to which we attribute function – both internally, with respect to the whole (e.g. the function of a heart in an organism), and externally with respect to the interaction between the whole and an environment (e.g. the ecological niche of the organism). For dynamical complexity, the number of interacting parts may be a useful proxy for the degree of complexity, whereas for a different context a measure of complexity could be the number of different relations, or functions, expressed through correlated parts and not the number of parts, *per se* (McShea, 2000). Another way we can think of complexity is the substitutability of parts for one another without affecting the behavior of the system. By this criterion, a Swiss watch is complex since one part cannot be substituted for another, whereas a gas composed of the same kinds of molecules would be simple since one gas molecule may be substituted for another gas molecule without affecting the behavior of the gas.

In the end, it is unlikely that complexity reduces to some necessary and sufficient essence. The term does not derive from a classical definition but more as a colloquial label for the experience of being cognitively overwhelmed (Andersson, Törnberg & Törnberg, 2014; Andersson & Törnberg, 2018) – which may happen in a number of ways. With human groups, complexity thereby relates to the extent to which cultural life places a cognitive burden on the group members. A group composed of students attending a lecture and a professor giving a lecture to the students would be a simple group for the group members since each member understands the roles and behaviors of the other group members. A group of politicians negotiating legislation, on the contrary, would be perceived by most as a complex group. We see politicians in shifting factions, each concealing additional layers of similar complexity in their own parties and networks, and each with different interests with regard to the legislation in question. This makes for a cognitively highly demanding environment where the distinction between the political game and its rules is blurred.

In these latter two examples, complexity relates to the number of different components that are involved with regard to what is required for the system to act in a coordinated manner and have consistent outcomes. This suggests that a first approximation to the complexity of a system is to measure complexity by the number of distinct system components that need to interact in a coordinated manner for the system to be coherent, resilient and to continue to operate as a sustainable system. By coherent is meant that the interacting components of the system do not, simply by their mode of interacting, lead to the collapse or breakdown of the system. Lack of coherency may arise for a variety of reasons and includes lack of fit between components, mismatch in the timing of the action of one component with another component, and so on. By resilience is meant the ability of the system to revert back to its normal operating state after an external shock or force impinges on the operating state of the system. By being sustainable is meant that the system can function in the same manner over repeated instances of the system's operation or over time scales much longer than the internal time scale of the system.

For human groups, an added consideration that needs to be considered is not only the performance of the system, but the support systems that may be required for the performance of the system in question. Thus, the system composed of a professor giving a lecture to a group of students is part of a larger system consisting of a university, the organization of university activities into events such as courses, an administrative system for the enrollment of students in courses, the assignment of professors to be the lecture for a class, the physical structure of the university, the position of a university in a larger community and so on. The complexity of a group, then, involves not just the phenomenal level of the behavior of the members of a group, but also the complexity at the ideational level of the cultural idea systems (Leaf & Read, 2012) that provide the shared framework within which group members interact.

This is why we cannot automatically assume that the complexity of an artifact, as an artifact, reflects the complexity of the cultural idea system underlying its production. A lower degree of artifact complexity does not, by itself, imply a lower overall complexity of the ideational system for which the artifacts are the instantiation. For example, hunter-gatherer societies face trade-offs between mobility and the production and transportation of artifacts that affect the complexity of the artifacts (Read, 2008).

We need, then, to take into consideration the difference between the complexity of an artifact as an artifact and the complexity of an artifact as the consequence of a cultural idea system that has conditioned the behavior of an artisan in producing the artifact. For the latter it follows that we need to consider how complexity relates to the modes, attributes or idiosyncratic features of an artifact that reflect the idea system and not just the number of parts. Consider how we might compare the complexity of a tool (a termite stick) made by a chimpanzee and a tool (Acheulean hand axe) made by a hominin ancestor (*Homo erectus*) of *Homo sapiens*.

Complexity of a Chimpanzee Termite Stick versus an Acheulean Hand axe

Termite sticks are made by chimpanzees to get termites out of termite mounds for consumption. The termite stick is made by breaking off a short branch from a bush or a small tree, then removing the leaves and any side subbranches from the branch, and then modifying the end of the tool with their teeth (see e.g., Sanz & Morgan, 2007, p. 430-431). This leaves a short stick that can be inserted into an opening in the termite mound. When the stick is removed from the termite mound, there may be termites attached to the stick that can then be licked off and eaten by the chimpanzee. In terms of number of parts, a termite stick has complexity C = 1.

Stone tools referred to as Acheulean hand axes were first made about 1.7 mya in Africa (Asfaw et al., 1992; Diez-Martin et al., 2015) by *Homo* ergaster/*erectus* and later by *Homo heidelbergensis* (Corbey et al., 2017) and, through the early expansion of *Homo* out of Africa, had a wide distribution across the Old World (Petraglia & Shipton, 2008). Their function, though, remains contentious. While there is general agreement that they were used for tasks involving cutting or scraping, other, very different, functions that have been suggested include throwing them as a weapon, or using them for social and/or sexual signaling (see references in Corbey et al., 2016). Early hand axes were made with a minimal amount of flake removal and later ones were well-made, showing the mastery of the technology of conchoidal flake removal. Hand axes continued to be made throughout the Lower Paleolithic for almost 1.5 my, with more recent hand axes dating to around 500 kya.

Hand axes have a characteristic flattened, generally symmetric shape with a shape ranging from lanceolate to ovate to orbiculate (Corbey et al., 2016). The shape is produced by removing flakes through flaking. In addition to making the shape of the hand axe by flake removal, the stone from which the hand axe was made was also made thinner by removing flakes from the front and back surfaces. For this reason, hand axes are said to be bi-facial. They vary in size from small to large. If the number of parts is the measure of complexity, a hand axe also has complexity C = 1, hence for this measure termite sticks and hand axes are equally complex.

One thing that C, the number of parts, does not measure is precisely the differences in the idea systems brought to bear in the production of a hand axe versus a termite stick. For C to be a relevant measure we must assume that the production processes behind the artifacts are of similar complexity.

The Acheulean hand axe brings us to the related question of whether, and if so to what degree, the complexity observed is cultural in the first place. If hand axes are a cultural object as understood by archaeologists such as Irving Rouse, this requires that they would have to have been produced in accord with a shared idea system. To some, the fact that hand axes had the same shape worldwide and for over 1,000 millennia suggests that their persistent common shape was due to genetic inheritance rather than to traditional inheritance, let alone to the inheritance of a shared system of ideas.

This judgement may be more intuitive than argued, of course. It seems, though, that cultural inheritance would not be capable of maintaining stasis in shape worldwide and for over a million years. Constraints could arise from different sources that are not mutually exclusive, including the stone flaking technology itself, the raw material used and the intended functionality of the hand axe. Another explanation is that the hand axe was a key *enabler* of a flexible system of adaptation and so was itself not subject to flexibility due to Generative Entrenchment (Andersson, Törnberg and Törnberg, 2014a; Wimsatt, 1999, 2015; Wimsatt & Griesemer, 2007). That is, altering its design may have had immediate cascading maladaptive repercussions in dependent parts of the cultural system. But, notably, the potential for strong cultural lock-in effects opens the door also for genetic scaffolding. If the hand axe came to serve such a critical role in hominin life, its faithful reproduction (a narrow reaction norm) may have been increasingly buttressed by genetic canalization of behavior through a Baldwin effect (Baldwin 1856; see also Sterelny, 2004). That is, genetic factors that made the desired outcome more likely would be selected for (Corbey et al., 2016; Foley, 1987; Richerson & Boyd, 2005).

But perhaps even more likely, and more preserving of the trademark flexibility of hominin behavior, is that cultural and genetic organization stabilized the transmission process itself. Something that was traditional, yet critically important, could moreover have acted as an evolutionary driver for derived general teaching and learning in *Homo* (Castro & Toro, 2014; Csibra & Gergely, 2009, 2011; Tehrani & Riede, 2008).

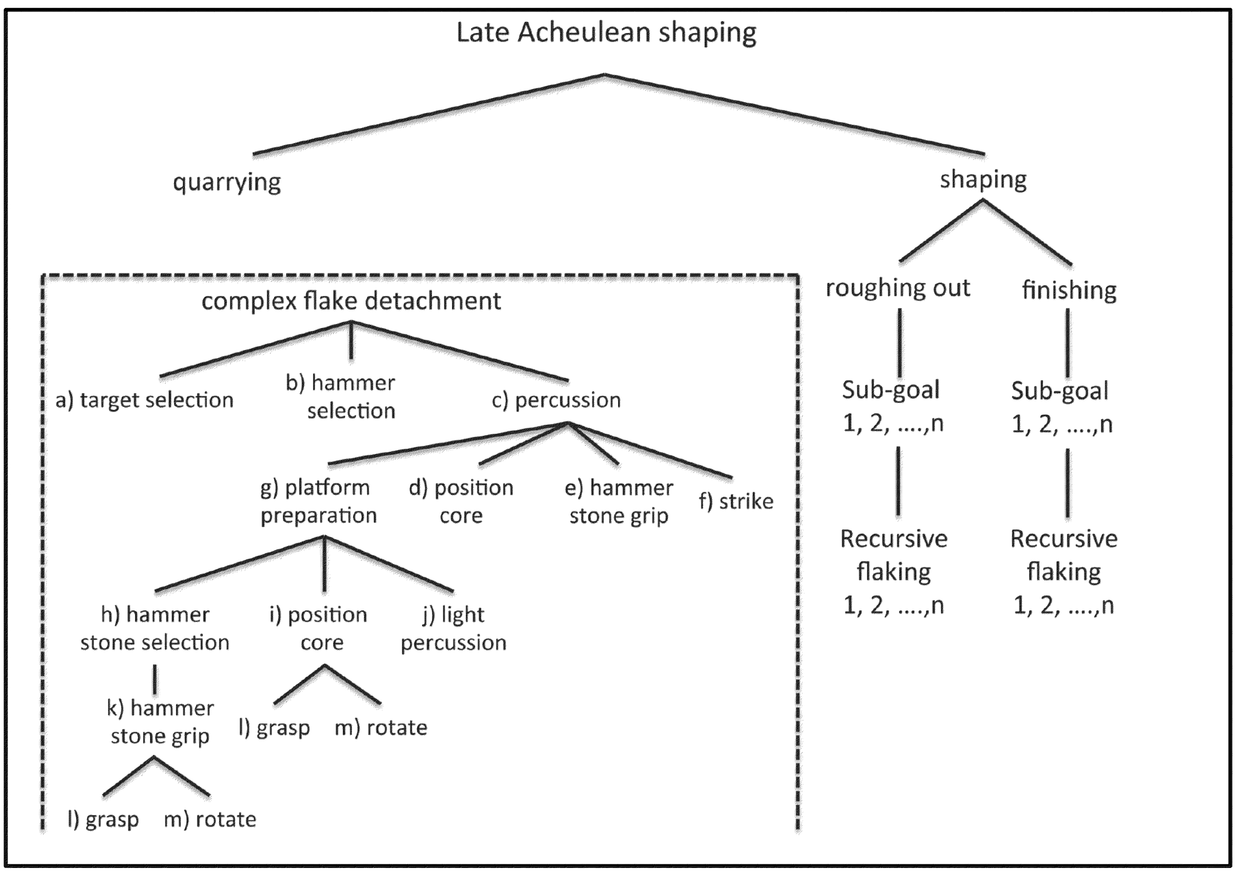


Figure 1: Handaxe subgoal hierarchy. Flake detachment: (a) select target, (b) select hammer, (c) select percussion. Percussion has sub goals: (d) position the core, (e) hammer stone grip, (f) strike core. Percussion also involves: (g) prepare platform, (h) hammer stone selection, (i) position the core, (j) use light percussion. Additional subgoals are: (k) hammer stone grip, (l) grasping, and (m) rotating the object. (Modified from Stout [2011, p. 1052] by Gärdenfors & Högberg [2017].)

Another constraint may have been the limited size of short term working memory (STWM) for the hominins making hand axes. STWM increased during hominin evolution from 2 ± 1 for the common ancestor of *Pan* and *Homo* (Read, 2006) to 7 ± 2 for modern *Homo sapiens* (Miller, 1956), so the STWM of *Homo erectus* would be about 4 – 5 (see Figure 2 in Read & van der Leeuw, 2008). Assuming the production of hand axes was at the upper bound of the cognitive abilities of *Homo erectus*, introduction of the more complex stone tool technology that marks the end of making hand axes would only occur after there was an increase in STWM. The stasis in *Homo* brain encephalization from around 1.5 mya to around 0,7 mya implies stasis in STWM over this time period, hence to stasis in stone tool technology as well.

Following Stout (2011), we will express an inferred idea system for the production of a hand axe as a hierarchical sequence of goals and sub-goals involved in the production of a hand axe. In Figure 1, the upper tree diagram identifies the major sub-goals involved in making a hand axe. This includes the hand axe shape that is produced by first roughing out the desired shape from a quarried nodule through flaking, and then making a more refined shape by the controlled removal of small flakes along the boundary of the hand axe and from the front and back surfaces. The boxed tree diagram in Figure 1 elaborates on the flaking process and shows that the complex flake detachment used in making hand axes involves several sub goals: the location on the object where a flake is to be removed, the choice of a knapping hammer, and the use of percussion as the means to remove the flake. The sub-goal of percussion removal has a series of sub goals and each of these has additional sub goals. Of these, some may be modes (assuming hand axes are cultural objects), in which case the sub goal that is a mode will be executed in a comparable manner by all artisans. Other sub goals refer to attributes which one artisan may accomplish in a

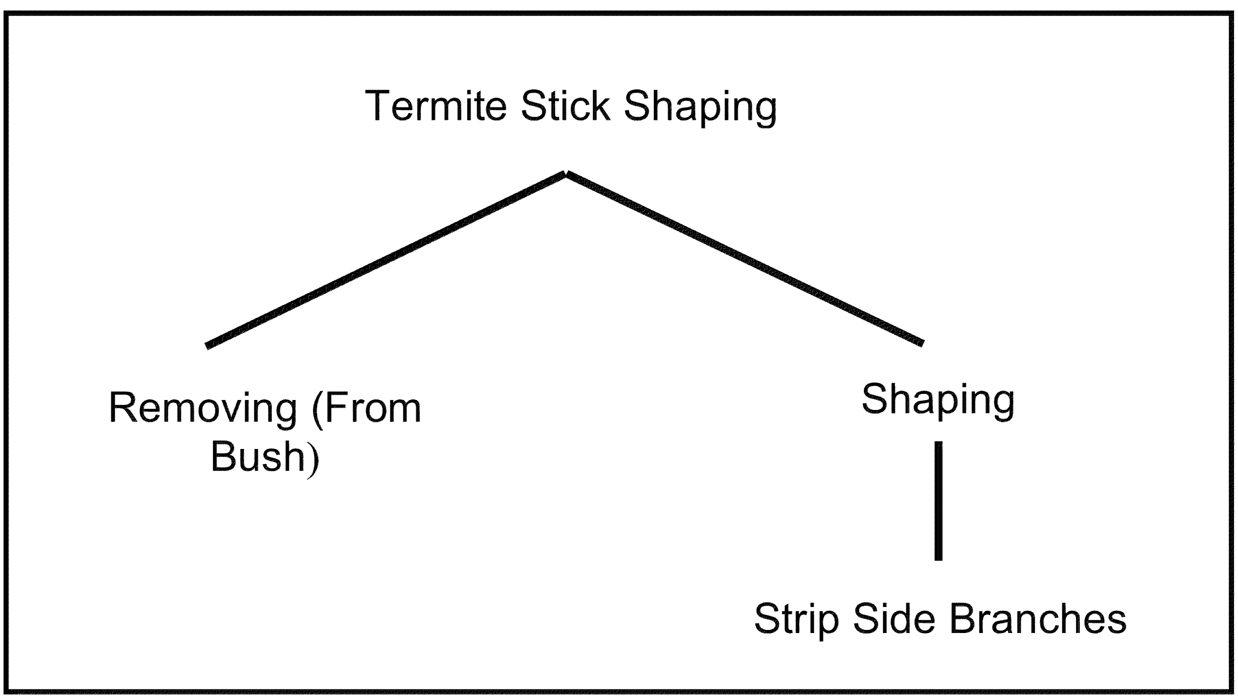


Figure 2: Subgoals for making a termite stick.

different manner than another artisan, and yet other subgoals may have idiosyncratic outcomes, such as (hypothetically) the orientation of a flake on one face of the hand axe in comparison to the orientation of a flake removal in the same position but on the other face of the hand axe. If we consider the sequence of subgoals involved in making a hand axe, its complexity is poorly measured by just the number of parts of the artifact.

We can compare the termite stick to the hand axe by working out a similar diagram identifying the subgoals involved in making a termite stick. First of all, none of the features of a termite stick are modes since chimpanzees cannot communicate with each other regarding features of a termite stick and then reach agreement on which features will be done in the same way by all chimpanzees making termite sticks. Some features, such as the diameter of the termite stick, are attributes in Rouse's vocabulary since a chimpanzee does not randomly break off a branch regardless of the diameter of the branch, but selects a branch having a diameter size that will fit into the openings of a termite mound. There will be similarity in diameter size across termite sticks since each chimpanzee needs essentially the same diameter size in order for the termite stick to fit into the termite mound. A feature such as length may be idiosyncratic since there is no single, optimal length, but a range of length values that make for effective termite sticks, hence a chimpanzee may not attempt to make any particular length within the range of feasible lengths for a termite stick. There is no community with a culture for making termite sticks conditioning the behavior of the chimpanzee. As indicated in Figure 2, the idea system tree implemented independently by each individual chimpanzee is shallow as only a few subgoals are involved. The initial subgoal is to remove a branch from a bush and then shape the branch. Shaping has the subgoal of stripping leaves and side branches from the branch, leaving a stick that becomes the termite stick.

Comparison of Figures 1 and 2 shows that from the perspective of the idea system involved in producing an artifact, an Acheulean hand axe is much more complex than a termite stick. Acheulean hand axes were used for processing multiple types of materials, which indicates that they were integrated into a wider cultural fabric of traditional practices. At the same time, some of the subgoals of the Acheulean hand axe appear to be constrained by the nature of the task; e.g., a hammer stone grip is used for holding the percussion stone used in flaking and so will be similar across artisans without requiring community consensus on how to grip a hammer stone, indicating that its production may contain some, but perhaps surprisingly little, and thereby robust, cultural scaffolding. The Acheulean hand axe appears to have been part of a holistic system that was inherited partly genetically and partly culturally.

This contrasts sharply with the termite sticks made by chimpanzees. If a chimpanzee community stops fishing for termites or "forgets" how to make termite sticks, they lose the termites but no other activities will be affected adversely by the loss. At most, if not a total lack, they have only a very tenuous system of interdependent activities.

We can refer to such decoupled traditions as statistically shared ideas. They are not culturally shared ideas. Even though individuals may statistically share the same idea, a statistically shared idea is not, by itself, a marker of a cultural attribute. This is especially clear with termite sticks. Each chimpanzee utilizes the same idea system shown in Figure 2 for the production of termite sticks, so chimpanzees statistically share the same idea system and termite sticks are all similar to each other. However, if they stop using termite sticks, no other practice will be affected in a major way since no other tradition relies significantly on termite sticks and, conversely, making termite sticks does not rely on other traditions, thus termite stick making does not have the coupling of traditions associated with cultural idea systems.

Complexity Through Implement Utilization

Another way that relating complexity just to measurable aspects of an artifact can be misleading regards the use of an artifact in a task. How the artifact is used to achieve the intended goal also involves an idea system. Consider the nut cracking by chimpanzees. This is a traditional practice that contains no artifacts with imposed forms at all, but instead selected objects are organized into an adapted system (that indeed itself may be viewed as an artifact). Complexity here resides in the ability to inter-operate all of the parts of this system, which is cognitively demanding to do. To crack a nut, a chimpanzee places a nut on a stone anvil and hits it with a stone, an action that cracks the nut shell and gives the chimpanzee access to the nut meat inside of the shell. The chimpanzee then eats the nut meat. Though a simple task for humans, it poses a challenge for some chimpanzees. The chimpanzees crack nuts in a group, so mature adults cracking and eating nuts provide models for juveniles to watch and to learn how to crack nuts. It takes about 2 years for infant chimpanzees to learn how to crack nuts by watching other chimpanzees crack nuts, beginning roughly at 3 years of age and only succeeding in cracking nuts when they are about 5 years of age, around the time a chimpanzee reaches physical and sexual maturity (Matsuzawa,

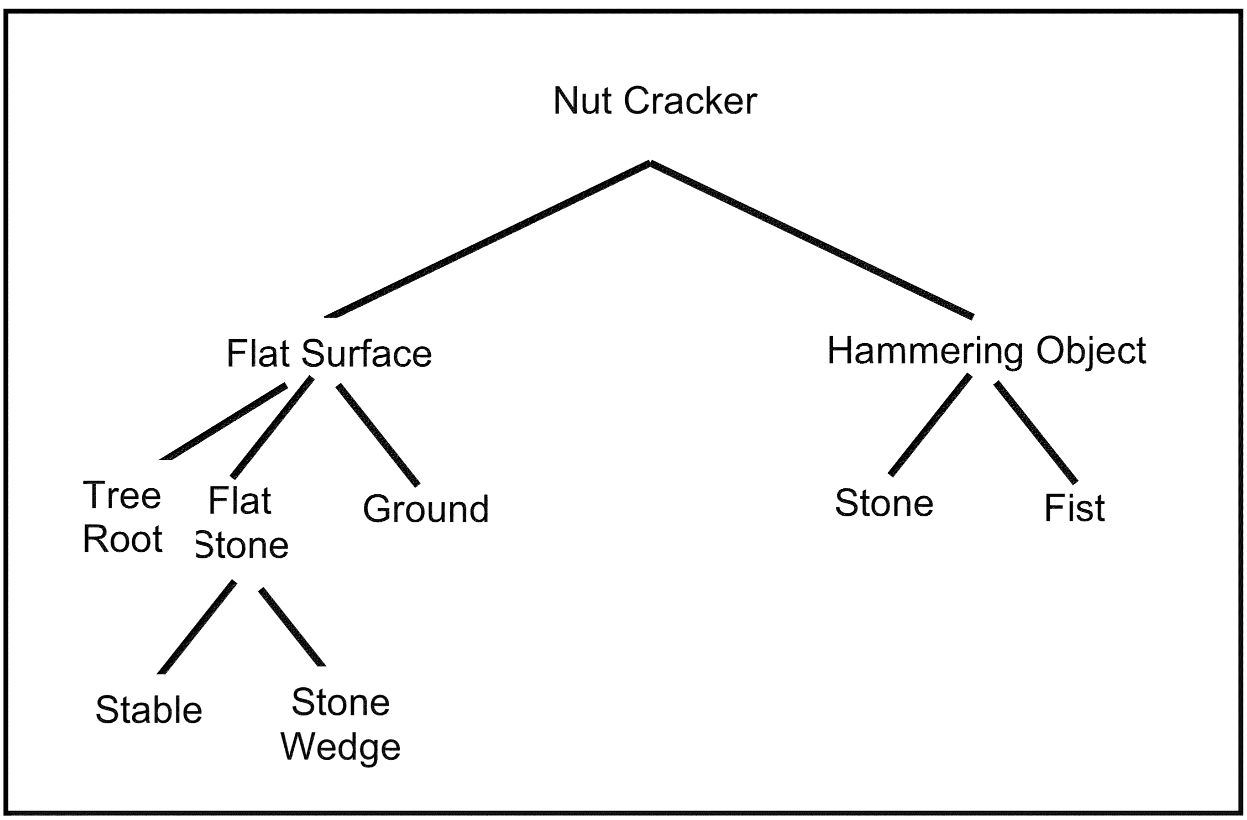


Figure 3: Subgoals for making an implement to crack nuts.

2007). Striking is the fact that about 20% of the chimpanzees never learn to crack nuts. Instead, these chimpanzees place a nut on the anvil and hit the nut with her/his fist, or they place the nut on the ground and hit the nut with a rock. They never learn the sequence: position a stone anvil on the ground, place a nut on the anvil, then hit the nut on the anvil with a stone.

We can divide this sequence into two parts: (1) production of the implement – an anvil and hammer stone – and (2) utilization of the implement. Figure 3 shows the system of ideas involved in making an implement for cracking nuts. Two primary subgoals are a flat surface and a hammering object. (That the hammering object is a subgoal can be seen in the fact that chimpanzees in Taï National Park in Côte D'Ivoire will collect and carry rocks with them that are then used to crack nuts [Boesch & Boesch, 1983].) The flat surface has three possible subgoals when subgoals are summed over different nut cracking chimpanzee groups: an exposed tree root with a flat surface (a subgoal for chimpanzees at Taï National Park [Boesch & Boesch, 1983]), a rock

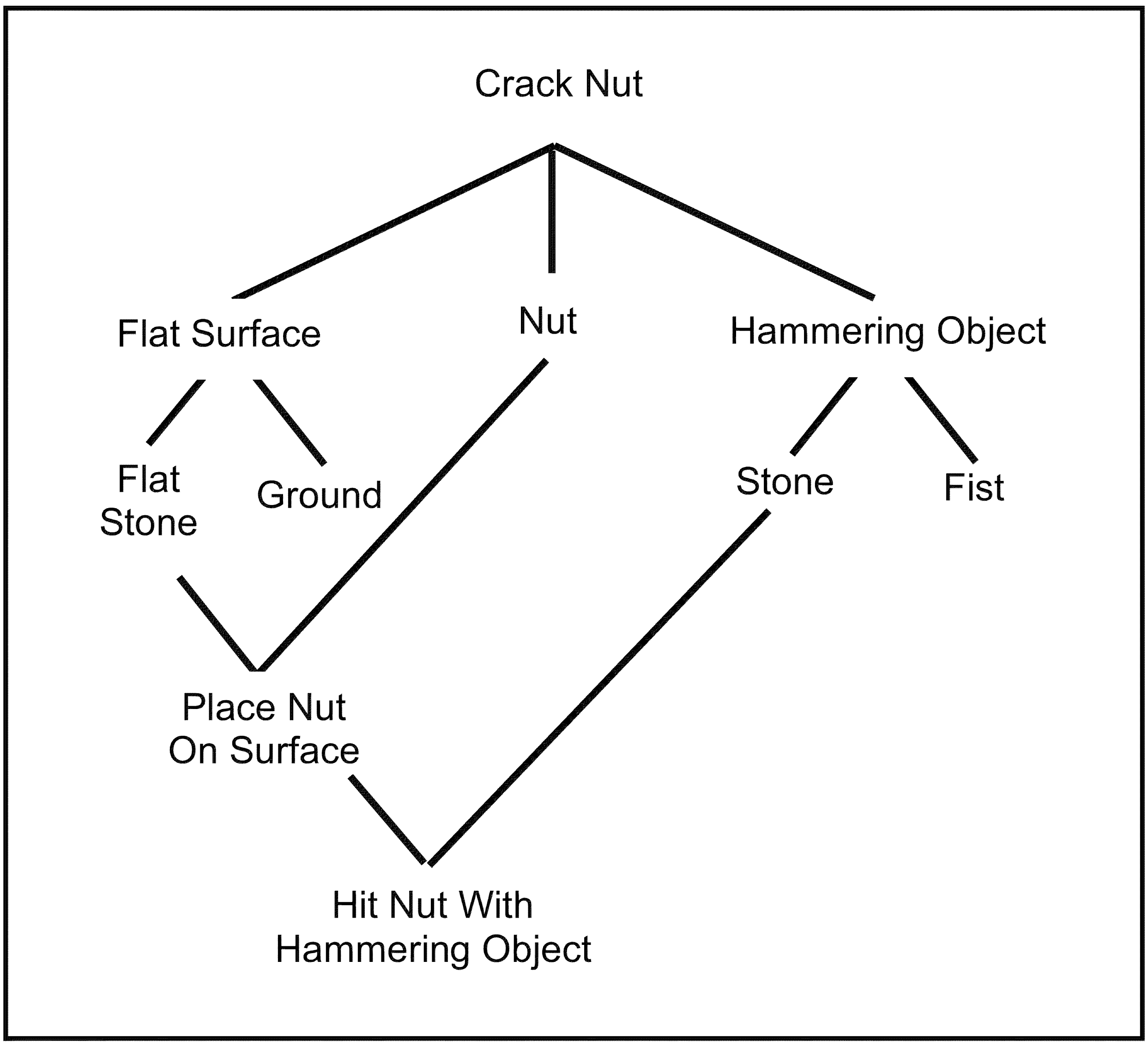


Figure 4: Subgoals for cracking nuts.

with one side relatively flat (a subgoal for chimpanzees at Bossou in Guinea [McGrew et al., 1997]), and the ground itself (this becomes a subgoal for some of the chimpanzees that never learn to crack nuts with an anvil). The flat stone has the subgoal of being a stable flat surface. The rock with a flat surface may already be stable. If not, chimpanzees have been observed to use small stones as ‘wedges’ to stabilize the rock (Matsuzawa, 1996). The hammering object has two subgoals: a stone that will be used for hammering a nut or, for some of the chimpanzees that do not learn to crack nuts, the hand will be made into a fist and the nut will be hit with the fist.

The nut cracking implement is used to crack nuts. Successful nut cracking involves a two-step hierarchy of actions: (1) place a nut on the anvil and (2) hit the nut with a rock hammering object. Successful implementation of the nut cracking device requires that three ideas be kept in mind simultaneously through STWM: use the anvil as the source of a flat surface, place a nut on the anvil and then strike the nut with a hammer stone. All three must be active in short term memory in order to implement correctly the hierarchical relationship of the two sub-actions: place nut on surface and hit nut with hammering object (see Figure 4).

Now consider the fact that 20% of the chimpanzees never learn to crack nuts even though they watch other chimpanzees successfully crack nuts. The unsuccessful chimpanzees either place the nut on the anvil and then hit the nut with a fist (Figure 5A) or place the nut on the ground and hit the nut with a hammerstone (Figure 5B). (Some chimpanzees possibly place a nut on the ground and then hit it with a fist, but this has not been reported.) In both cases, the chimpanzee makes use of a natural object – the ground or a fist – hence the action only requires STWM = 2.

While there is a strong correlation between the size of STWM and measures of cognitive abilities, the relationship is not completely deterministic. Cognitive abilities involve a variety of

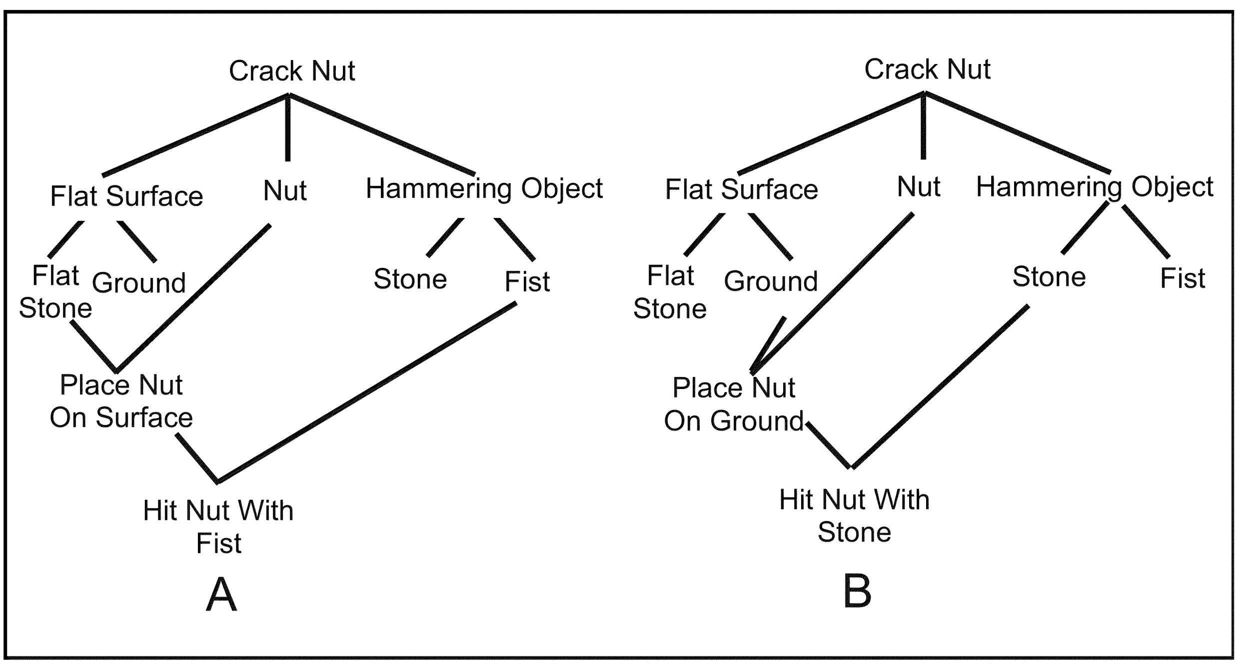


Figure 5: Unsuccessful idea systems. (A) Hit nut with fist. (B) Place nut on ground.

neurological processes. An overly simplistic model for the performance of chimpanzees measured by the ability to learn to crack nuts would be that chimpanzees with a STWM = 1 or 2 do not learn to crack nuts and it is the chimpanzees with STWM = 3 that learn to crack nuts. A more realistic model would be that chimpanzees with STWM = 1 never learn to crack nuts, some chimpanzees with STWM = 2 do not learn to crack nuts and other chimpanzees with STWM = 2 do learn to crack nuts through using other neurological processes that "compensate" for the limitations of STWM = 2, and all chimpanzees with STWM = 3 learn to crack nuts. Let us use this characterization of nut cracking by chimpanzees to form a thought experiment in which we include the role of phenotypic transmission for a non-genetic trait such as cracking nuts.

Thought Experiment

We can make a thought experiment based on chimpanzee nut cracking as a way to explore the validity of the treadmill model that purportedly links cultural evolution of cultural complexity through the DIT model to the size of the population of interacting individuals. The purpose of the thought experiment is to incorporate the assumptions of the treadmill model and then to determine whether the connection it claims to have demonstrated between the interaction population size and cultural complexity is valid under the conditions assumed for the treadmill model.

The essence of the treadmill model is its dependence on three non-controversial assumptions, though how the third assumption is implemented is problematic. First, it is assumed that phenotypic transmission through imitation is subject to transmission degradation. Second, it is assumed that skillful individuals, even when imitating a degraded target, can produce a non-degraded version of the degraded target. In effect, this assumption only requires that when imitating a degraded target, a skillful individual can recognize in what way just imitation of the target will lead to a degraded execution of the target and can then correct the degradation. In effect, the assumption is based on the idea that a skillful person does not merely imitate a target in a rote manner, but through emulation (Tomasello, 1996) by recognizing, in the case of nut cracking, how effective nut cracking must proceed and whether simply imitating the target is effective or not, and if not, what would need to be modified to make nut cracking effective. The third assumption follows directly from the fact that the expected value, N, for the number of persons of a

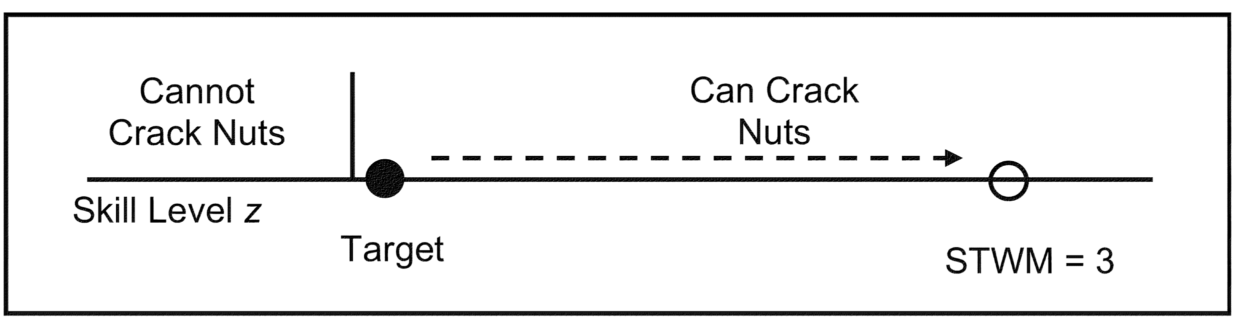


Figure 6: The variable z measures the skill level of a chimpanzee. The vertical line shows the minimum skill level needed to crack nuts. Skillful chimpanzee with STWM = 3 imitates a target doing nut cracking with a low skill level (solid disc). Dashed arrow shows that the skillful chimpanzee, through emulation, cracks nuts with a high level of skill (open disc).

given skill level in a population is given by $N = p \times n$, where p is the probability of a person having the specified skill level and n is the population size. It follows that the expected number of individuals with a specified skill level increases with population size. Next, construct the thought experiment.

For purposes of the thought experiment, assume 20% of the chimpanzees have STWM = 1, 60% have STWM = 2, and 20% have STWM = 3. Assume that chimpanzees with STWM = 3 are skillful, so according to the second assumption they can learn to crack nuts not in just a rote manner, but are able to put together a mental nut cracking model that identifies for them what makes nut cracking effective, such as the surface of the anvil should be horizontal and not sloped, the hammerstone should be roughly spherical in shape, the size of the hammerstone should be such that it is easily held in the hand and is large enough so that the momentum imparted from the arm by a striking motion suffices to crack the nut but not so great as to smash completely the nut. Consequently, we can model the imitation of a target by a chimpanzee with STWM = 3 as shown in Figure 6. Even if the skill level for the nut cracking device employed by the target for cracking nuts is low (e.g., the anvil is at an angle or the hammerstone is too small), a chimpanzee with STWM = 3 and imitating the target will end up putting together a nut cracking device that can be used skillfully to crack nuts.

For a chimpanzee with STWM = 2, the skill level achieved, according to assumption 1, will be less than that of the target (see Figure 7). If the target skill level is close to the boundary for the skill level required for being able to crack nuts, then the implement produced through imitation will not suffice for cracking nuts (not shown). Assuming the degree of skill level lost through imitation is a function of the imitation process itself, it follows that the mean skill level of those with STWM = 2 will decrease in each round of nut cracking. For a chimpanzee with STWM = 1, assume the skill level attained, regardless of the skill level of the target, is below the skill require for nut cracking.

Assuming that the most skilled individual will be the target individual for phenotypic transmission through imitation (as in Henrich, 2004), it follows that if there are individuals with STWM = 3, then the average skill level in the population will not decrease since individuals with STWM = 3 perform at a high skill level even after imitation. Individuals with STWM = 2 will

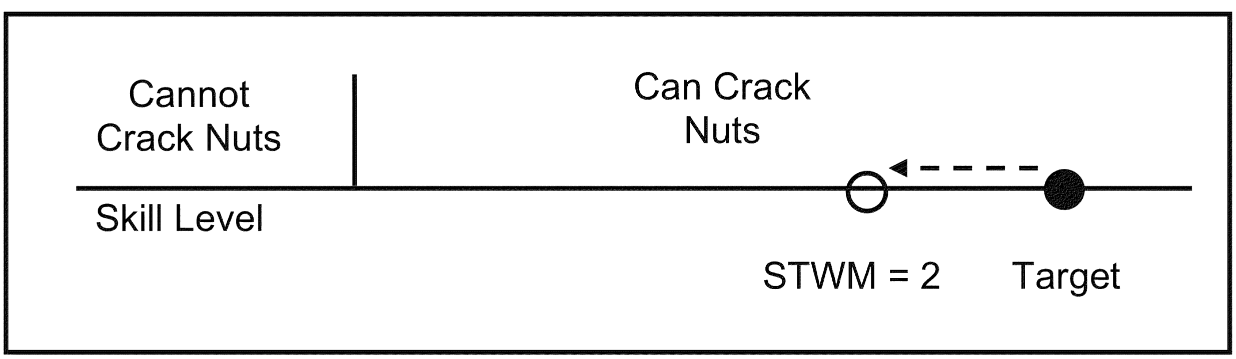


Figure 7: Chimpanzee with STWM = 2 (open disc) imitates a target doing nut cracking with a high skill level (solid disc). Dashed arrow shows that the chimpanzee, through imitation, cracks nuts with a lowered level of skill.

perform at a degraded skill level after imitation, but some of them will still able to crack nuts at a degraded skill level. However, their degraded skill level will not decrease further so long as, in the next generation, there are STWM = 3 individuals who crack nuts effectively and so they can imitate individuals with STWM = 3 in the next generation of nut cracking. Individuals with STWM = 1 cannot crack nuts and have a skill level below the level required for cracking nuts even when the target they imitate has a high skill level for cracking nuts.

If, for whatever reason, individuals with STWM = 3 are lost from the population, then the average skill level will reduce with each generation of imitation since the resetting of the target skill level by individuals with STWM = 3 no longer occurs and the STWM = 2 chimpanzees imitate at a skill level below that of the target. Since the target for the next generation will, at best, be an individual with STWM = 2 and will have a degraded skill level, the average skill level for the STWM = 2 individuals will decrease from one generation to the next. Chimpanzees with STWM = 1 have skill levels after imitation below the target skill level and below what is required to crack nuts, regardless of the skill level of the target. Eventually, then, the population will lose the ability to crack nuts.

Conversely, for a population that initially does not have the skill required to crack nuts, and if, for whatever reason, an individual with skill level STWM = 3 is introduced into the population, this individual may work out nut cracking and if so, will become the target individual. The skill level of the STWM = 2 (and STWM = 1) individuals will be reset when the individual with STWM = 3 becomes the target individual, so individuals with STWM = 2 will now imitate the new target individual and be able to crack nuts, though in a degraded manner as indicated in Figure 7 since they do not imitate the target individual perfectly. The average skill level of the implements will increase.

The thought experiment, then, performs in the same manner as the treadmill model (see Henrich, 2004) – a trait requiring a high skill level for its effective occurrence may be lost if highly skilled target individuals are no longer present, perhaps, according to the treadmill model, by a sufficient decrease in the interaction population size so that it is unlikely that there will be a highly skilled person in the reduced population size. Conversely, if the presence of a highly skilled individual becomes more likely due to increase in the interaction population size, and if this highly skilled person invents (or reinvents) the trait and becomes the target individual, then the trait will increase in its frequency of occurrence, though in degraded form except by individuals who are highly skilled. Thus, the thought experiment makes the same predictions as the treadmill model.

At first glance, the thought experiment seems to verify the treadmill model. However, closer examination shows that this is not the case. The treadmill model assumes that the addition of a highly skilled person into the population is due to increase in the interaction population size by virtue of Assumption 3, namely that in smaller populations the expected number of extremely skilled persons is smaller than in larger populations. Consider in more detail the assumption of the treadmill model that a more highly skilled individual is introduced into the population by increase in the interaction population size. For illustrative purposes, consider how this relates to the occurrence of individuals with high IQ scores under the usual assumption that $\mu = 100$ and = 15 for IQ scores. For chimpanzee nut cracking, we assumed in the thought experiment that 20% of the population is skilled at nut cracking (STWM = 3). The top 20% of IQ scores would corre-

	Table 1: Expected Number of Persons								
Expected # of persons	Probability								
	.0125%	0.000625	0.00125	0.03125	0.0625	0.125	0.5	1	IQ ≥ 155
	2%	0.1	0.2	1	10	20	80	160	IQ ≥ 130
	20%	1	2	10	100	200	800	1600	IQ ≥ 113
Population size		5	10	50	500	1000	4000	8000	

spond to those individuals with IQ ≥ 113. From Table 1, in a population with 5 persons, E[N | n = 5, IQ ≥ 113] = 1 ("expected number N of persons with IQ ≥ 113 in a group of n = 5 persons is 1 person"), so even with a family-size group there should be at least one individual who is reasonably skilled. For IQ ≥ 130, the IQ score for a person to be considered mentally gifted, a group of size 50 – slightly larger than a residence group in a hunter-gatherer society – would have E[N | n = 50, IQ ≥ 130] = 1. For a hunter-gatherer society of n = 500 persons, we would expect 10 persons with IQ ≥ 130. For a population of size n = 8,000, which is the estimated interaction population size of Tasmania before isolation from the mainland (Henrich 2004), and with p = 0.0125% (corresponding to an IQ ≥ 155, which is close to the highest score of 160 in the Stanford-Binet IQ test) we would have E[N | n = 8,000, IQ ≥ 155] = 1. For an interacting population of size n = 4,000, the estimated interaction population size of Tasmania after it was isolated from the mainland due to the increase in ocean levels, we would have E[N | n = 4,000, IQ ≥ 155] = 0.5; thus, for groups of size n = 4,000, on average one-half of them would not have any individual with IQ ≥ 155. This indicates that it would require a skill level of at least IQ = 155 in order for the reduction of the interaction population size from n = 8,000 to n = 4,000 to result in the smaller population possibly not having any individual with an IQ matching the highest IQ level that could be found in at least one individual in an interacting population with n = 8,000). In other words, the demographic effect upon which the treadmill model depends, namely that a larger population may have a highly skilled person but a smaller population may lack an equally skilled individual, only applies to genius-level skills when considering populations the size of simple hunter-gatherer societies. Yet it does not require a genius to make a simple bone point and simple clothing. If it did, this would require that the far more complex tools found in other hunter-gatherer societies would require even substantially higher IQs than this for their invention.

These data show clearly that it is only with skill levels at the gifted-to-genius range, and only for groups smaller than 50 persons that are likely to lack a skilled person. For hunter-gatherer societies, with a modal value of n ~ 500-600 persons (Read, 2012a), we expect 10 persons to be in the 98th percentile or higher (see Table 1). We conclude, then, that the likelihood of not having a highly skilled person in a population due to the population size only applies to small populations with at most 10 – 50 persons, and even then, it only applies to highly skilled persons. Thus, the treadmill, as a driver for increasing the skill level with which a task is performed, is applicable at best to small sub- populations within a hunter-gatherer society and does not apply meaningfully to the entire hunter-gatherer society. Nonetheless, the proponents of the treadmill model claim that it is supported by experimental data and by data for hunter-gatherer societies. Consider the experimental data first.

Experimental Data for Testing the Treadmill Model

A variety of experiments have been performed, purporting to support the treadmill model applied to hunter-gatherer societies (e.g., Caldwell & Millen, 2010; Derex et al., 2013; Kempe & Mesoudi, 2014; Muthukrishna et al., 2014). Common to all of these experiments is the small size of the group involved in the experiment. For example, Caldwell & Millen used groups with n =1 to n = 3 individuals and their experiment did not show any difference in performance between the two group sizes. Their experiment was critiqued by Muthukrishna et al. as involving a single imitation model too simple in its design (a paper airplane) to show any effect on performance by the small difference in group size in the experiment.

These small group sizes, alone, invalidate an experiment as a test of the treadmill model since the effects determined on small population sizes, whether positive or negative, cannot be assumed to scale up to population sizes comparable to the size of hunter-gatherer populations. Further, the experiments, for the most part, are not designed to test the effect of the interaction population size on group performance due to a small group not having a highly skilled individual, while a much larger interacting population is likely to have a highly skilled individual, as posited in the treadmill model. For example, the independently performed experiments by Murhukrishna et al. and by Kempe & Mesoudi each used several imitation models to avoid the problems with the Caldwell & Millen experiment, but still used small group sizes. Murhukrishna et al. compared the performance of a group with size n = 1 with a group with size n = 5, while Kempe & Millen used a group size of n = 3 for the larger group in the comparison. Though the larger groups performed better, it is not clear whether this is due to group size difference leading to the inclusion of a more skilled person in the group, or to some other factor.

A general problem with these experiments, then, in addition to the very small group sizes, is that even when different population sizes are part of the experiment, there is no determination of the skill levels of the individuals in a small group versus a larger group, hence any difference in performance between a smaller and a larger group cannot be attributed to the larger group having a more skilled target person by virtue of it being a larger group, which is the core argument of the treadmill model. For example, the experiment by Derex et al., which considers different group sizes (n = 4, 8, 16) and does find differences in performance according to group size. However, the differences are in the wrong direction. Performance by larger groups in their experiment is worse than the performance by smaller groups (Andersson & Read, 2014). In addition, the only other effect they found is simply the well-known fact that smaller populations are more prone to drift (Andersson & Read, 2014). In reply to these observations, Derex et al. (2014, p. E2) state: "Even if explained by sample size effect, this supports the group size hypothesis: *sample size effect is expected to be the main mechanism by which group size affects cultural evolution*" (emphasis added). However, the claim of the treadmill model is not the well-known effect of drift as a function of sample size, but the consequences of a larger group more likely including a highly skilled individual. None of the experiments purporting to support the treadmill model, then, actually test the demographic process modeled in the treadmill model (Andersson & Read, 2016).

Data on Hunter-gather and Fishing Societies

Data on tool-complexity in hunter-gatherer and fishing societies are unequivocal. There is no correlation between either population size or the interaction population size and the complexity of artifacts in any of these societies. To see this, consider four sets of data. The first set is the data on Tasmanian hunter-gatherers since they were considered to provide an ideal data set for testing the treadmill model due to archaeological documentation of the disappearance of bone points in the Tasmanian archaeological record around 8,000 BP when Tasmania was isolated from the Australian mainland after a global rise in ocean levels. This, it has been argued, lead to the maladaptive loss of warm clothing (Henrich, 2004, 2006). The disappearance of bone points in the archaeological record, coupled with the presumed reduction in the size of the interacting population prior to the isolation of Tasmania, seemed to provide unequivocal evidence for the treadmill model. The second set will be meta data sets consisting of ethnographic data on the population size and population density of hunter-gatherer societies that leads to statistically testing whether the predicted correlation between the interaction population size and tool complexity is verified. The third set is composed of the complex tools made by Inuit groups in the Arctic. The fourth is a data set on Oceania fishing groups for which it is claimed (Kline & Boyd, 2010) that the complexity of fishing hooks correlates with the population size, as predicted by the treadmill model.

Tasmania Data Set

The bone points found in several archaeological sites on Tasmania have been referred to as complex tools and the isolation of Tasmania has been interpreted as causing a change from an interacting population with around 8,000 persons when there was a land connection between Tasmania and the mainland to a population of about 4,000 persons when the rising ocean levels isolated Tasmania from the mainland (Henrich, 2004). Thus, it appears that the loss of the supposedly complex bone points occurred following the reduction of the size of the interacting population on Tasmania and consequently, according to the treadmill model, the likely skill level of the most skilled person in Tasmania was reduced due to the decrease in the interaction population size. As discussed above, though, this would require a skill level corresponding to an IQ ≥ 155 for making bone points for this to happen, hence bone points would have to be complex tools. However, bone points are simple tools (Mulvaney & Kamminga, 1999); see also Figure 1in Read, 2012b).

Figure 8 shows the idea system for making bone points. As can be seen by comparing Figure 8 with Figure 2, the complexity of the idea system for bone points is essentially the same as that of the termite sticks made by chimpanzees, yet *Homo erectus*, with a cranial capacity about 2/3 of the cranial capacity of modern *Homo sapiens*, regularly made Acheulean hand axes that involved a far more complex idea system (see Figure 3) even though there is no evidence that *Homo erectus* individuals would have been part of interacting populations larger than 4,000 individuals. Thus, it would not have required an interacting population of 8,000 individuals for there to be individuals with the skill needed to make simple bone points.

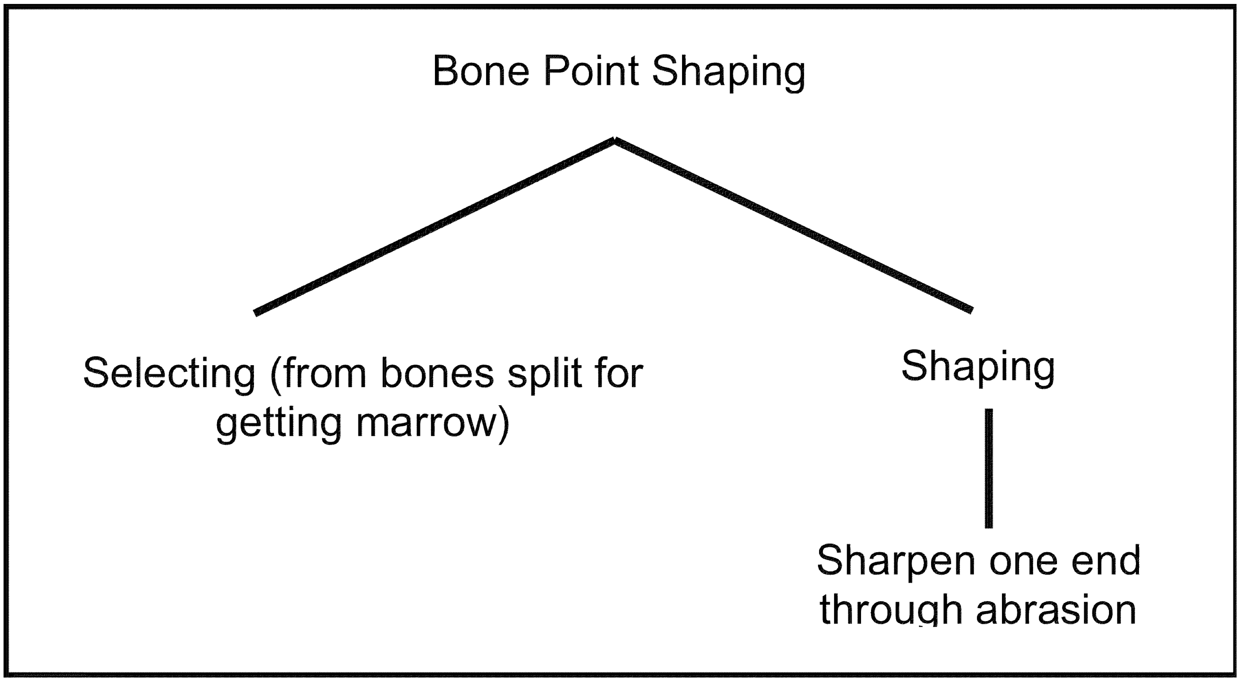


Figure 8: Subgoals for bone point shaping.

Despite claims to the contrary (e.g., Henrich 2004), the Tasmanians did not lose the ability to make clothing. The first European explorers to contact the Tasmanians noted that the Tasmanians, especially the women, wore simple cloaks when the temperature was especially cold. Prehistorically, during the ice ages, when the Tasmanians made clothing using bone points as awls, they were not making complex clothing like the Inuit, as claimed by Henrich (2004, 2006), but were making simple clothing (Gilligan, 2014) that matched their thermal need for clothing given their degree of biological adaptation to a cold climate (Gilligan, 2007). The loss of bone points, which coincided with world-wide warmer weather and the increase in ocean levels (see Figure 2 in Read, 2012b), is most parsimoniously explained by the Tasmanians no longer needing the warmer clothing that required bone points for its manufacture and so they stopped making bone points (Gilligan, 2014).

Meta data sets on Hunter-Gatherers

Eight publications (Buchanan, O'Brien & Collard, 2015; Collard, Buchanan, & O'Brien, 2013; Collard et al., 2005, 2011b, 2013a; Read, 2006, 2008, 2012b) have compared the population size of hunter-gatherer societies to the complexity of their implements. Seven of these publications have found that there is no correlation between population size and implement complexity and one publication found a negative correlation. One other publication (Collard et al., 2011a) found no statistically significant correlation between risk and population size, but this is just a statistical artifact. The researchers only included cases from the same region, hence the cases included in the analysis had a narrow range of risk values and, as is well-known in statistics, constraining the variance in the predictor variable to a small range of values will always reduce the magnitude of the correlation between the predictor variable and the outcome variable in a regression analysis. Altogether, none of the statistical modeling in these eight publications support the treadmill model. Another four publications have used published data to argue that there are conceptual problems with the treadmill model (Andersson & Read, 2016; Collard et al., 2013b; Read, 2011; Vaesen et al., 2016).

The proponents of the treadmill model have countered by saying that the sample of hunter-gatherer populations used to compare population size with implement complexity are skewed towards hunter-gatherer groups from Northwest America and all of these publications have used the census population size rather than the interaction population size (see, e.g., Henrich, 2006; Henrich et al., 2016; Kline & Boyd, 2010). The first objection is not valid since the treadmill model is not region specific and predicts a high correlation between the interaction population size and the complexity of tools regardless of the region where the hunter-gatherer society is located. The second objection is technically correct but the conclusion they make is not correct. Henrich et al. assume that because the treadmill model posits a causal relationship between the interaction population size and tool complexity, but not between the census population size and cultural complexity, no correlation is expected between the census population size and tool complexity. However, their conclusion requires that the census population size varies randomly with respect to the interaction population size, for if there is any positive correlation between the two measures (regardless of causation), either both correlate with tool complexity or neither correlates with tool complexity. Though Henrich et al. discuss problems with the accuracy of census data for hunter-gatherer groups in detail, they provide no data showing that the census population size varies randomly with respect to the interaction population size. Further, their argument ignores the fact that the interaction population size *is* a function of the population density. Hence the predicted correlation between the interaction population size and tool complexity may be tested by computing the correlation between tool complexity and population density. When this comparison is made, the correlation between the population density and tool complexity is found to be zero (Read, 2006). In sum, the predictions from the treadmill model are falsified empirically by meta data sets on hunter-gatherer groups.

Inuit of the Arctic Region

Data on the Inuit of the Arctic region and the problems these data pose for the treadmill model have been discussed extensively in Read (2012b). Briefly, the Inuit made some of the most complex tools of any hunter-gatherer group. Within these data, the Greenland Inuit data are especially problematic for the treadmill model due to the Angmaksalik Inuit of eastern Greenland making a harpoon with 33 parts, the most complex implement of any hunter-gatherer group (Oswalt, 1976). The Angmaksalik Inuit had a population size of 420 persons in the earliest census of Greenland (Petersen, 1984, Table 2), an order of magnitude smaller than the population of Tasmania that was said by Henrich (2004, 2006) to be too small to make simple bone points and clothing. With regard to the maximum possible interaction population size of the Angmaksalik Inuit, the total population of Inuit in Greenland was about 6,000 persons, hence even if all the Inuit in Greenland interacted with one another, despite groups being separated by 100's of kilometers of rugged coastline, the interaction population size was at most about the same size as the interaction population of Tasmania. More realistically, the Angmaksalik Inuit would had contact at most with the Inuit on the southern coast of Greenland, but the fact that their mtDNA haplotype frequencies distinguish them from the Inuit of southern Greenland (Helgason, 2006) implies that drift was a more important factor than migration in structuring their mtDNA haplotype frequencies: "the current differences indicate that drift has outweighed gene flow" (Helgason, 2010). Thus, there was little interaction between the east and southern coast, let alone between Inuit on the east coast and the west coast of Greenland. Yet even if we assume the interaction population included both the east and the south coast Inuit, then we still have the striking contradiction that even through there were supposedly no target persons in Tasmania sufficiently skilled to make simple bone points and clothing, the Angmaksalik Inuit made a vastly more complex implement – a harpoon with 33 parts – with at most the same interaction population size. No counterargument to the conclusions drawn from these Inuit data has yet been published.

Oceania Fishing Hooks

In their study of the complexity of fishing hooks made by groups on the Oceania Islands with a subsistence economy, Kline and Boyd (2010) claim that the complexity of their fishing hooks varies positively with the interaction population size. However, close examination of their data and analysis, discussed in Read (2012b), shows, for five reasons, that this is not the case.

First, Kline and Boyd state, correctly, that each group in their data set must have the same economic basis, namely a subsistence economy (Kline, 2010). However, Hawaii, one of the

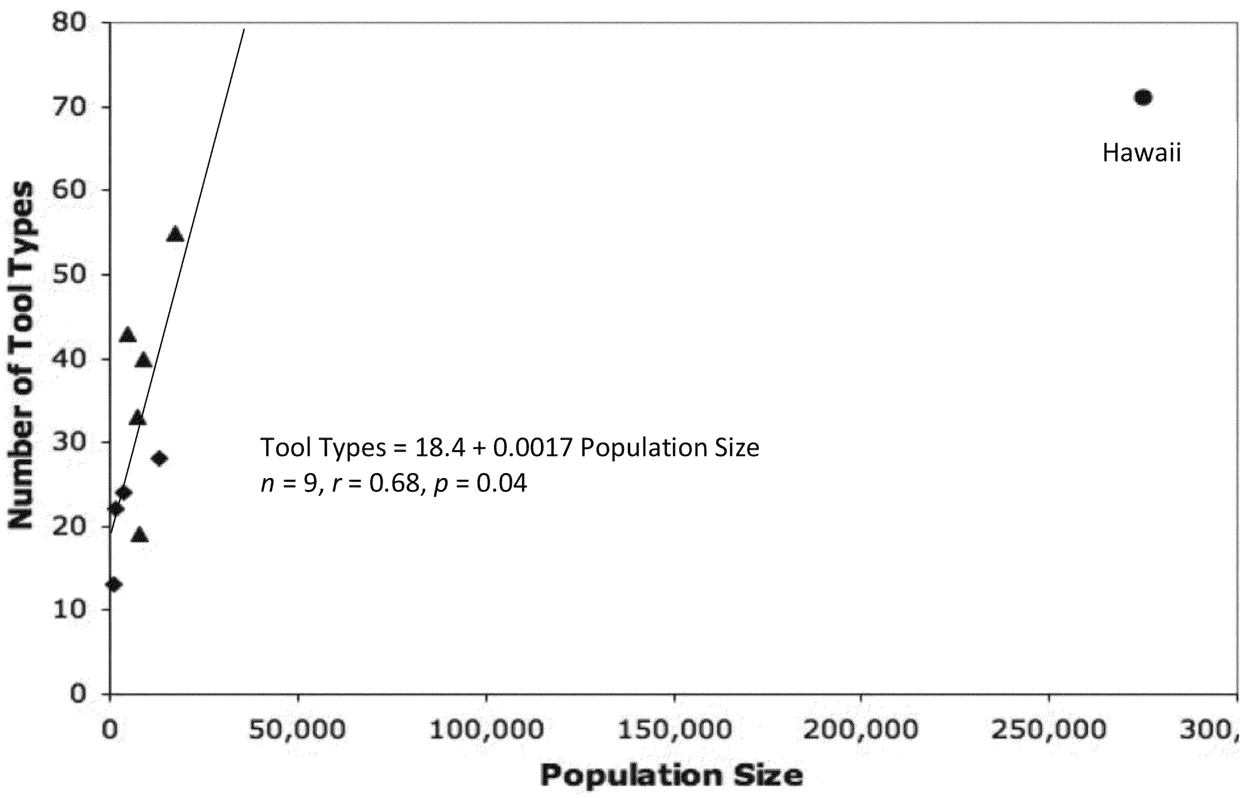


Figure 9: Linear relationship between population size and number of tool types for 9 Oceania Islands (excluding Hawaii). Triangles: groups with high rate of contact with other groups. Diamonds: groups with a low rate of contact with other groups. Solid disc: Hawaii. Hawaii is an outlier.

groups in their sample, had a barter, not a subsistence, economy, and had extensive fish farms characterized as being comparable "to integrated farming systems developed in ancient China and Egypt" (Costa-Peirce, 1985, p. 328). When Hawaii is removed from the data set, the correlation between population size and tool complexity is not significant at the 5% level (n = 9, r = 0.61, p = 0.08) (Read, 2012b). Next, in order to assess the relationship between the interaction population size and the complexity of fish hooks, Kline and Boyd (2010) measured whether groups that had a high contact rate with other groups and groups that had a low contact rate with other groups were non-randomly distributed around the regression line computed for population size used as a predictor of tool complexity. The test for a random distribution was not significant even at the 60% level (p = 0.64, Fisher Exact Test), hence groups with a high contact rate did not have more complex tools than groups with a low contact rate, contrary to the prediction from the Treadmill model (Read, 2012b).

Second, they note that the treadmill model implies that the relationship between tool complexity, measured by number of tool types, and population size should be a concave curve, but is, in fact, linear, as can be seen visually in Figure 9. It is clear from Figure 9 that Hawaii is a statistical outlier, thus, Hawaii violates not only their criterion of only including groups with a subsistence economy but is statistically an outlier in comparison to the linear relationship between population size and number of tool types that characterizes the other n = 9 groups in their sample.

Third, their measure of risk is the frequency of unusual weather events such as typhoons, whereas the risk of concern in the risk hypothesis is the risk of failing to be successful on a food

Table 2: Tool Complexity, Contact Between Groups and Ocean Currents

Island Group	**Mean TU**	**Population Size**	**Contact**	**Ocean Currents**	**(Ocean Currents) * (Population Size)**
Malekula	3.2	1,100	1	1	1,100
Chuuk	3.8	9,200	2	1	9,200
Santa Cruz	4.0	3,600	1	2	7,200
Trobriand	4.0	8,000	2	2	16,000
Tikopia	4.7	1,500	1	1	1,500
Yap	5.0	4,791	2	2	9,542
Lau Fiji	5.0	7,400	2	3	22,200
Tonga	5.4	17,500	2	2	35,000
Manus	6.6	13,000	1	3	39,000

TU: Number of Technical Units. TU is defined as "an integrated, physically distinct and unique structural configuration that contributes to the form of a finished artefact" (Oswalt 1976:38)
Contact: Frequency of contact with other groups, 1 = Low, 2 = High
Ocean Currents: 1 = Protected, 2 = Partially Protected, 3 = Not Protected
Data on TUs, Population Size and Contact are from Kline and Boyd (2010)
Data on ocean currents are from Read (2012b: Appendix)

procurement episode, not the risk associated with occasional extreme weather events. Their failure to show a correlation between the frequency of typhoons and fish hook complexity is not a valid test of the risk hypothesis.

Fourth, their statistical analysis shows, for the Oceania fishing data, that it is population size and not the interaction population size that is a predictor of tool complexity, contrary to Henrich (2004, 2006). In their regression analysis, the variables used for predicting tool complexity include (1) population size and (2) interaction population size. However, once population size is included in their regression model, interaction population size fails to be included as well (Kline & Boyd, 2010). Hence, contrary to the treadmill model, their analysis implies that population size, not the interaction population size, is the predictor of tool complexity. However, even this result is misleading since Kline and Boyd did not include a relevant risk variable in their analysis.

Fifth, and most critical, the risk of fishing as a daily subsistence activity is primarily related to whether an island is protected from daily ocean currents by coral reefs or small islands, not the frequency of occasional events such as typhoons, the risk measure used by Kline and Boyd. The risk imposed by ocean currents to the success of a fishing episode may be measured indirectly by the extent to which an island is exposed to ocean currents (see Table 2). When this measure of risk is included in the regression analysis, the only variable significantly predicting tool complexity (measured, following Kline and Boyd [2010], by the mean TU value for fishing

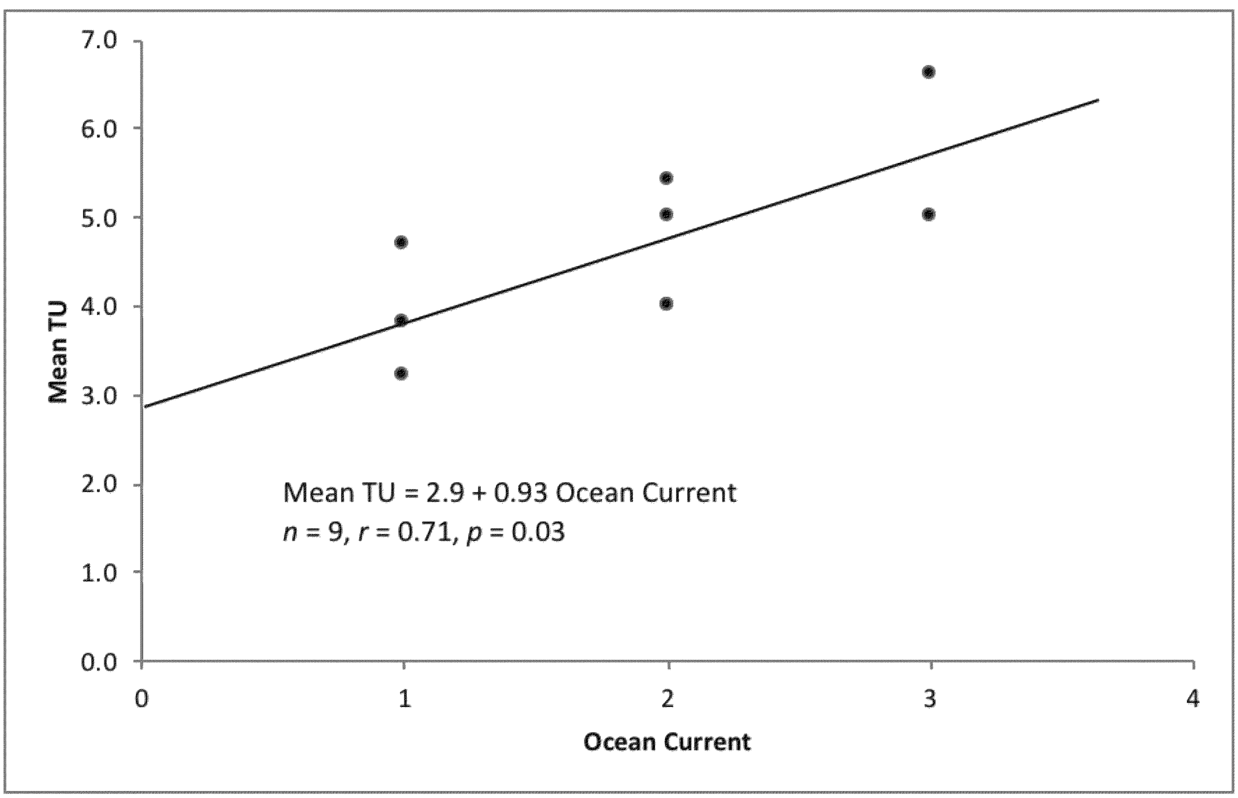
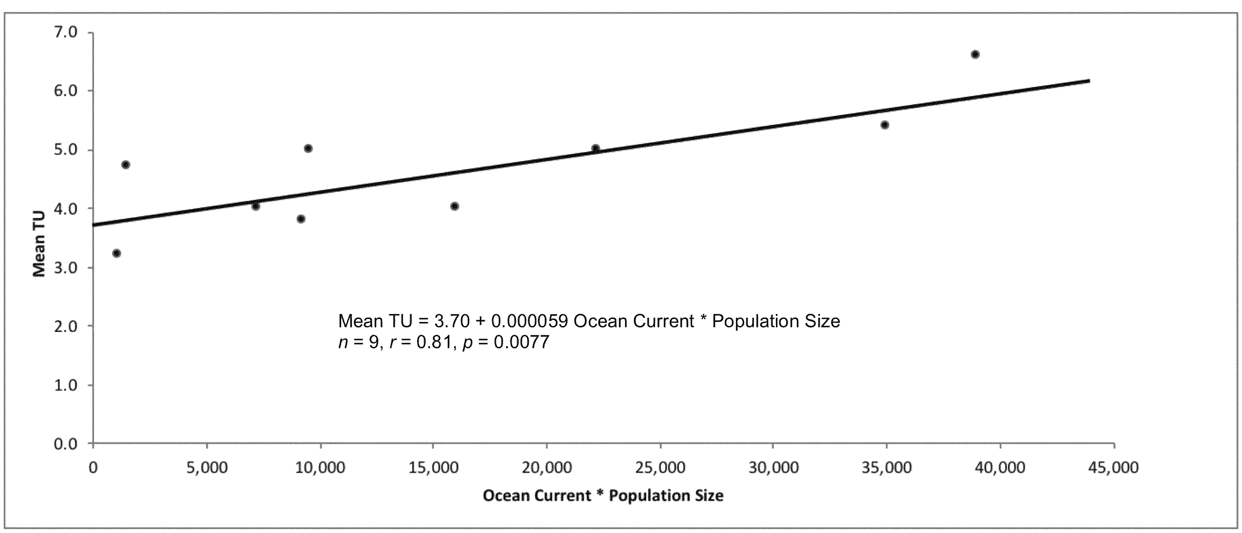


Figure 10: Linear regression model between Mean TU and Ocean Current is significant at the 5% level.

gear for each group) is the degree of protection from ocean currents (see Figure 10). The correlation, $r = 0.71$, between Ocean Current and TU is significant at the 5% level ($p = 0.03$) (Read, 2012b). Thus, their data show that the fishing hooks vary in complexity according to fishing risk and not according to interaction population size or to population size. No counterargument has yet been published with regard to the conclusions derived in Read (2012b) and summarized here.

In addition, but not discussed in Read (2012b), although the correlation between population size and tool complexity ($r = 0.61$), is not significant at the 5% level ($p = 0.08$), that tool complexity may increase with population size cannot be ruled out since the power of the statistical test is not high due to the small sample size ($n = 9$). These data, though, are unlike the meta hunter-gatherer data where the correlation between population size and tool complexity is numerically almost identical to zero. This is not the case here. It is possible, then, given the low power of the statistical test, that population size is relevant to tool complexity for the two large groups, Tonga and Manus, each with a population size an order of magnitude greater than the population sizes for the other groups. Both groups are in a higher risk environment, as measured by exposure to ocean currents, and each has greater tool complexity than is the case for the other Oceanic groups (see Table 2). Further, as with the meta data, there is an interaction effect (discussed in Read, 2008) between the measure of risk, Ocean Current, and population size that predicts tool complexity. The correlation between Mean TU and the interaction effect, (Ocean Current) * (Population Size), is $r = 0.81$, which is significant at the 1% level ($p = 0.0077$) (see Figure 11). This result is contrary to the treadmill model as it does not involve the likelihood of a more skillful person being present in a larger population in comparison to a smaller population, but only that negative effects from a high-risk environment are more pronounced with a larger population than a smaller one and the response to this interaction effect is to make more complex implements in order to compensate for the combination of a risky environment and a large population size (compare Santa Cruz with Tonga and Lao Fiji with Manus in Table 2), as predicted by the risk hypothesis.



Risk Hypothesis

The risk hypothesis (Torrence, 1983, 1989, 2000) relating artifact design complexity by hunter-gatherers to risk in food resource procurement has been widely cited and accepted in the archaeology literature. The hypothesis is based on the fact that food resource procurement carries risk in the sense that any episode of food procurement may be unsuccessful, for a variety of reasons, ranging from factors over which individuals have some control to factors over which they have little or no control. Three of the factors that affect risk and are central to the risk hypothesis are, first, the probability that a food procurement episode will be successful; second, the number of possible procurement episodes in each yearly cycle; and third, the quantity of food procured when a procurement episode is successful. These factors determine whether the quantity of food resources that can be obtained matches the need of a group for food resources. Of these three factors, the maximum number of possible food resource procurement episodes is largely outside of

Figure 11: Regression model between Mean TU and the interaction effect, (Ocean Current) * (Population Size). The regression model is significant at the 1% level.

the control of hunter-gatherers, the quantity obtained is partially under their control through choices made regarding which resource to pursue, and the probability of a successful food procurement episode once a resource is identified can be modified directly by hunter-gatherers through the design of the implements they make. The risk hypothesis focuses on this last factor.

The maximum number of possible procurement episodes is (to a first approximation) proportional to the length of the growing season since the growing season demarcates the period of time over which food resources are relatively abundant versus when they are less abundant and so are harder to procure. The length of the growing season varies from 365 days a year in equatorial areas to a few days a year in the extreme Arctic. The probability of a successful food procurement episode can be increased by hunter-gatherers in a number of ways, ranging from the knowledge they have regarding the yearly pattern for the location of food resources to actions taken by hunter-gatherers to ensure that any attempt to procure a food resource is successful, especially when procuring mobile resources through hunting and/or fishing. It is here where the complexity of implements used in the procurement of food resources comes into play.

Any implement used to obtain an animal resource depends upon completing subtasks such as how close the hunter must get to the animal for an implement to effectively wound, kill or disable the animal being hunted, the means for transferring energy from the hunter to the implement, the accuracy with which an implement can strike or hit the hunted animal, the design of that part of the implement which will penetrate into an animal or otherwise disable the animal, and so on. For implements composed of a single part, all of the subtasks involved in procuring an animal through the use of an implement must be done by the implement as a whole, hence a compromise will be necessary in the design of that implement since the design effectiveness with which each of these subtasks is carried out cannot be maximized simultaneously. Increasing the number of parts, with each part designed for better accomplishing a subtask, makes it possible to optimize each part of the implement for the subtask it will carry out.

That implements will be made with more parts (hence, will be more complex when complexity is equated with number of parts) in order to reduce risk through increasing the probability that a procurement episode is successful is virtually self-evident. There is a cost, of course, when making an implement with more parts. The likelihood that this cost is counter-balanced by the increased effectiveness of an implement made with several parts should be inversely proportional to the length of the growing season and proportional to the size of the resource that can be obtained since the risk of not obtaining enough food resources increases as the growing season gets

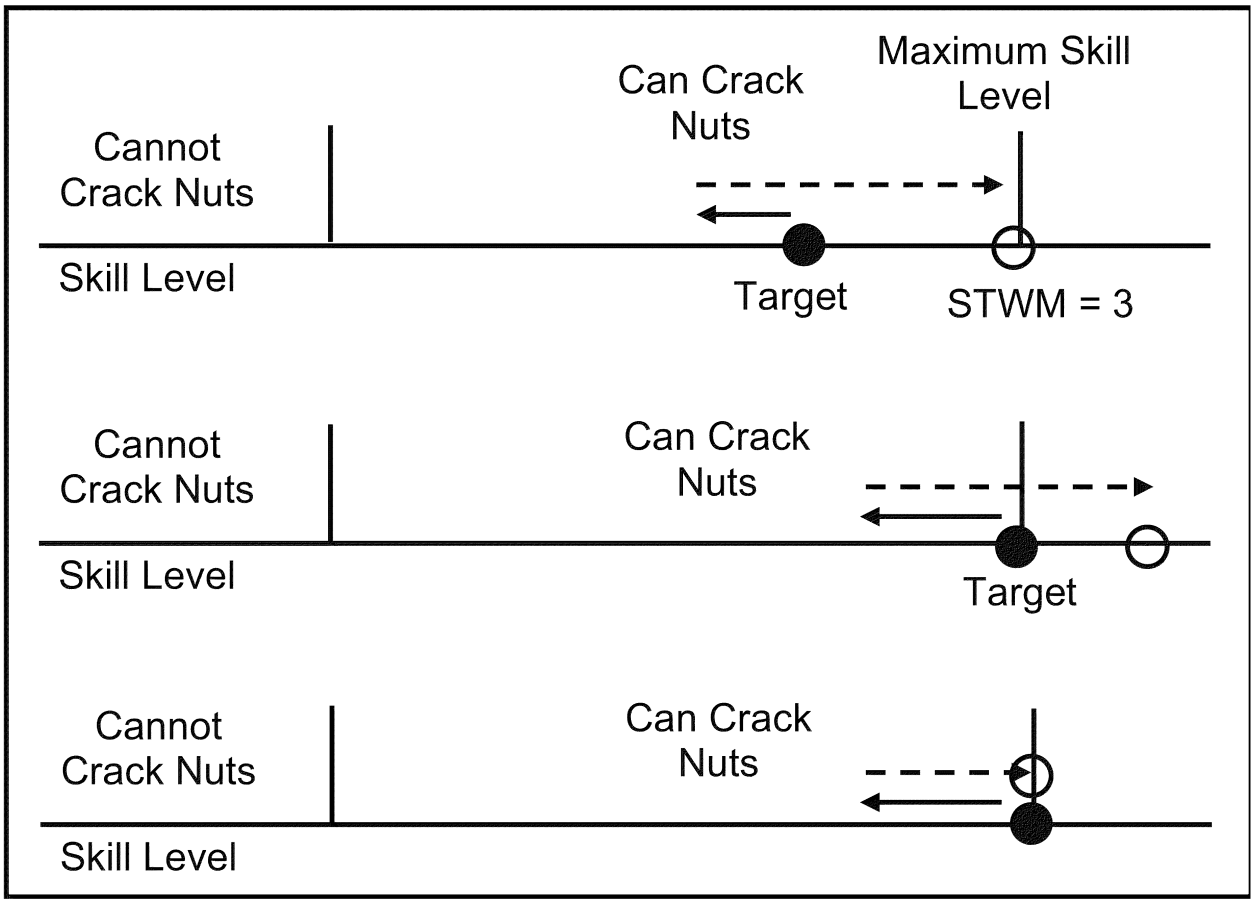


Figure 12: **Top**: Chimpanzee with STWM = 3 (open disc) imitates target (solid disc) nut cracking with a moderate skill level. Solid arrow shows degradation of imitated skill level due to imitation. Dashed arrow shows increase in skill level of imitator through emulation. **Middle**: Chimpanzee with STWM = 3 (open disc) imitates target from previous round of imitation nut cracking with a high level of skill. Solid arrow shows degradation of imitated skill level due to imitation. Dashed arrow shows increase in skill level of imitator through emulation under the assumption that increase in skill due to emulation is independent of the target skill level (Henrich 2004). The chimpanzee with STWM = 3 achieves a new, higher skill level under this assumption. **Bottom**: Chimpanzee with STWM = 3 (open disc) imitates target from previous round of imitation nut cracking with a high level of skill. Solid arrow shows degradation of imitated skill level due to imitation. Dashed arrow shows increase in skill level of imitator through emulation under the assumption that increase in skill due to emulation is inversely proportional to the target skill level. The chimpanzee with STWM = 3 does not achieve a new, higher skill level under this assumption.

shorter. This leads to the risk hypothesis with its claim that the complexity of implements should track risk, where risk may be measured by, for example, the length of the growing season, or by some other environmental measure relating to risk. The risk hypothesis was tested successfully by Torrence (1989) using latitude as a proxy measure for risk since climate differences affecting the growing season are associated with latitude.

Contrary to the risk hypothesis, the treadmill model hypothesizes a demographic constraint on increase in the complexity of implements derived both from the statistical property that the expected number of individuals in a population with a given skill level increases with the population size and from an invalid assumption used to translate this statistical property into a driver for tool complexity. The treadmill model assumes two parts to phenotype transmission through imitation: first, degradation of the skill level of the target due to imperfect imitation (see solid arrows in Figure 12) and second, innovation that compensates for the degradation (see dashed arrows in Figure 12). If the compensation is greater than the degradation, the net result would be an increase in the skill level over what is produced through imitation (see Figure 12, middle). For this to occur, it must be assumed that the increase in skill level through innovation is essentially independent of the skill level of the target. According to this assumption, if the target is an implement produced by the most skilled individual, then the imitator, through innovation, will produce an implement with skill level greater than the current maximum skill level in the population and the average skill level in the population would increase. The assumption that the degree of increase in the skill level through innovation is independent of the target skill level is incorrect. It would imply that if the imitator imitated the most skilled implement that he or she can produce, then he or she will now produce an implement requiring a higher skill level, contradicting the assumption that the target item is the most skillfully crafted implement that the imitator can make. Instead, the increase in skill level through innovation decreases with the increase in the skill level of the target, and the increase will, at most, be close to the amount of degradation for a target produced by the most skilled individual (see Figure 12, bottom).

The risk hypothesis assumes that most, if not all, the implements made by hunter-gatherers can be made with skill levels easily found in populations the size of a hunter-gatherer residence group (see Table 1). Even the most complex implements do not require skill levels beyond, say, the 98th percentile of individual skill levels, and there will be individuals with this skill level in a hunter-gatherer group. Yet the treadmill model assumes the opposite. The treadmill model hypothesizes that increased complexity is limited by the interaction population size and even the task of making simple clothing in Tasmania with simple points could only be done effectively by the highest skill level found in a population of 8,000 interacting individuals and this skill level would not likely be found in a group of 4,000 interacting individuals. According to the treadmill model, introducing more complex implements or tasks would require increasing the size of the interaction population sufficiently so that there would now be an individual with the skills needed to make an even more complex implements, while at the same time an individual with these skills would not likely be found in the previous, smaller interaction population.

The contrast between the implications of the treadmill hypothesis and the risk hypothesis for hunter-gatherer groups is stark. The treadmill hypothesis implies that we should find a clear and strong relationship between the size of the interaction population and the complexity of implements, whereas the risk hypothesis implies that there should be no or little relationship between the interaction population size and complexity. Further, the treadmill model is silent on the relationship between implement complexity and risk. Instead, any correlation between risk and implement complexity would require that there must be selection for a larger interaction population as the means for relating complexity to risk. Thus, the treadmill model implies the structural model: risk → interaction population size → implement complexity. Hence the treadmill model implies that the relationship between risk and implement complexity must be attenuated by the interaction population size.

However, a detailed analysis (see Read, 2008) of the relationship between risk (measured by the number of growing months), implement complexity, frequency of group movement by a hunter-gatherer group (also posited to be inversely related to implement complexity; see Shott, 1986) and the collector versus forager strategy for resource procurement proposed by Binford (1980) – and still widely used by archaeologists (Sutton, 2000) – does not show an attenuated effect between risk and implement complexity, but instead shows precisely the opposite. The

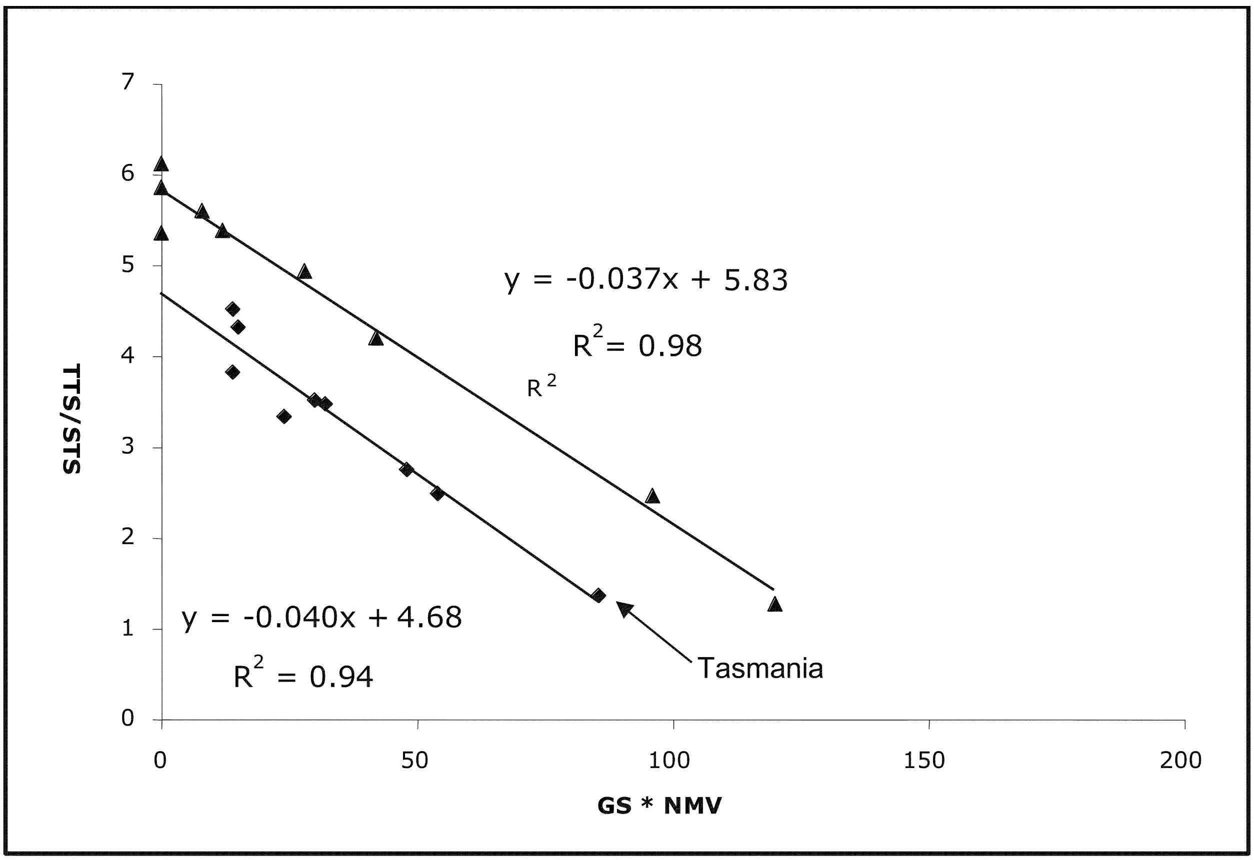


Figure 13: Upper Regression Line: Linear model for the collector strategy. Triangles (from left to right): Angmaksalik Inuit, Inglulik Inuit, Tareumiut Inuit, Tanaina, Ingalik, Twana, Nabesna, Ingura, Tiwi. Lower Regression Line: Linear model for the forager strategy. Squares (from left to right): Owens Valley Paiute, Copper Inuit, Tlingit, Nharo, Klamath, Caribou Inuit, Chenchu, Surprise Valley Paiute, Tasmania. GS: Growth Season, NMV: number of annual moves; TTS: Total Number of Technounits -- elaborateness; STS: Total Number of Subsistants -- diversity. STS, TTS data are from Oswalt 1976; GS, NMV data from Binford 2001: Tables 4.01, 4.07, 5.01.

analysis yields a regression model (see Read, 2008 for details) that accounts for about 96% of the variability in implement complexity (see Figure 13). In Read's analysis, the hunter-gatherer groups are the same as those used by Collard (2005) and Read (2006) (with the exception that the Andamese and Aranda hunter-gatherer groups are not included as they are clear outliers; see top part of Figure 5 in Read, 2008), and complexity is measured by TTS/STS, the average number of parts per implement, which is the measure for implement complexity used by other researchers.

Striking in Figure 13 is the statistical split in the data set into two distinct clusters, each with a linear fit for the regression of implement complexity (TTS/STS) on the interaction effect between length of growing season and mobility (GS * NMV). The upper cluster in Figure 13 is composed of 9 hunter-gatherer groups, 6 with a collector strategy, and the lower cluster is composed of 9 hunter-gatherer groups, 6 with a foraging strategy. Of the 6 out of 18 mismatches between strategy and cluster, the Chenchu are tropical foragers who were becoming incipient agriculturalists (Binford, 1980) and classified as collectors for this reason. Of the 5 remaining groups, only one is not questionably included in the wrong group (Read, 2008). With the Chenchu correctly coded as foragers, the null hypothesis that the groups are distributed around the regression lines independently of their classification as foragers or collectors may be rejected at the $\alpha = 5\%$ significance level since $P(x \geq 13) = 0.048$ for a binomial distribution with 18 trials and probability $p = 0.5$ of a success, x.

The two clusters in Figure 13 are consistent with Binford's (1980) observation that a foraging strategy uses mobility to track seasonal and spatial variation in resource availability by mapping people to resources, whereas a collector strategy averages out seasonal and geographic variation in resource availability by mapping resources in a region to people. A foraging strategy, then, tends to be less resource intensive than a collector strategy since population growth is constrained by Liebig's Law of the Minimum for resource density for a forager strategy, whereas population growth is constrained by the average resource density for a collector strategy. This expected difference in the strategies is supported by the demographic changes that occurred in the Southern Plateau in Western America after there was a shift from a forager to a collector strategy for obtaining food resources. Starting about 3,500 B.P., there was a rapid shift in the Southern Plateau from a mobile, forager strategy to a collector strategy that was then followed by an exponential rate of population growth until about 3,000 B.P. (Chatters, 1995).

A collector strategy, by averaging over spatial and temporal variability in food resources, can maintain a higher population density, but since this also depends on more intensive resource exploitation as population density increases, it follows, according to the risk hypothesis, that implement assemblages will be more complex for collector hunter-gatherer groups under the same conditions as a forager hunter-gatherer group. As can be seen in Figure 13, for a given value of GS * NMV, groups with a collector strategy have more complex tools, as predicted from the risk hypothesis but not from the treadmill model. According to the treadmill model, more complex implements and modes of resource procurement will correspond to larger interaction population sizes, hence to higher population densities since population density is a proxy for the interaction population size.

With regard to claims that Tasmania had unusually simple implements for a hunter-gather group (e.g., Diamond, 1978; Jones,1977; Oswalt, 1976), it can be seen visually in Figure 13 that

Tasmania, despite its simple tools for resource exploitation, is not out of line with other hunter-gatherer groups but is on the regression line for a forager strategy. The simplicity of their implements and the loss of bone points, though used by Henrich (2004) as an index of a maladaptive change driven by an externally imposed reduction of their interaction population size, can be more plausibly accounted for by their mobility in combination with the length of the growing season in Tasmania, and by no longer needing bone points to make simple garments as protection against the cold when the last ice age ended with global warming in the early Holocene, beginning from 11,700 BP.

That 96% of the variability in the complexity of implements is accounted for by the regression model implies that either the treadmill model must account for the risk model or, if it is independent of the risk model, it accounts for at most 4% of the variability in the complexity of implements. Since the interaction population size does not correlate with implement complexity and the interaction population size is the driving factor in the treadmill model, it follows that the two models are independent and so the treadmill model accounts for, at most, 4% of the variability in the complexity of implements for this data set, thus relegating the treadmill model to a virtually non-existent role when modeling the causal basis of implement complexity in hunter-gatherer societies.

Conclusion

In this article we have reviewed some of the issues stemming from current models regarding the drivers of cultural complexity and cultural evolution. Our concern has not been with the form of the models but with assumptions inherent to their implementation, especially for the small-scale, hunter-gatherer societies that typified the evolution of *Homo sapiens* prior to the Holocene. In particular, we have taken issue with the implementation of what is colloquially referred to as a treadmill model in which a demographic factor, the interaction population size, is asserted to be the driver of cultural complexity through the DIT model of evolution. The DIT model integrates genotypic and phenotypic transmission of traits, with the former central to evolution in the biological domain and the latter to evolution in the cultural domain. However, cultural evolution, as it is implemented in the DIT model, depends on defining cultural traits by reference to just the transmission part of Tylor's seminal definition of culture as a "complex whole." In so doing, the DIT model redefines traditions as cultural traits, thereby losing sight of the way that what constitutes culture involves far more than just the mode of transmission. We claim that DIT thereby becomes the proverbial streetlight under which a search is conducted for something that likely lies outside the area of illumination.

Culture has to do with shared idea systems and their transmission. Traditions and their transmission are, of course, of central importance for understanding culture, but it is only part of the issue, and a major drawback of DIT is that it is predisposed to reducing culture to a collection of traditions. Culture, in the full sense of the term, means that cultural evolution operates at the organizational and not at the population level assumed in the DIT model of evolution (Lane et al., 2009).

The DIT simplification of culture also predisposes us to accept the convenient idea that artifacts may be taken as the substance of cultural evolution. This runs counter to the arguments of archaeologists regarding material culture as the *instantiated consequence* of cultural idea systems. Culture has to do directly with shared idea systems, and only indirectly with artifacts as the instantiation of those shared idea systems (Read, 2007).

For artifacts, cultural complexity needs to be related to the cultural idea systems that guide their production and use, and not just to the intrinsic complexity of tools measured, for example, by counting the number of parts. The latter, by itself, leads to anomalies such as the conclusion that a termite stick made by a chimpanzee and a hand axe made by *Homo erectus* are equally complex since both are implements consisting of a single part. The number of parts does relate to complexity when considering how implements are made with multiple parts so that each part can be optimized for a particular function in the use of that implement. Thus, a termite stick and a hand axe are equally complex with regard to the way each is designed to do a task without subdividing the implement into parts designed to do a single function within the overall task. At a different analytical level, when we consider the idea systems involved in the production of a hand axe in comparison to that of a termite stick, the far greater complexity of the idea system involved in making of a hand axe in comparison to making a termite stick comes to the fore (compare Figures 2 and 3).

The treadmill model has been very influential in cultural evolutionary thinking about cultural complexity and its causes. It attempts, through a mathematical model, to reduce all of this, using the framework of the DIT model of evolution, to the statistical fact that the highest skill level found in a large population is likely to be greater than the highest skill level found in a small population. The model ignores the other statistical observation that a skill level likely to be found in a large population and simultaneously not likely to be found in a small population (that is, the skill level needed to satisfy the condition posited by the treadmill model for an increase in average skill level) requires that the skill level be in at least the $(1\text{-}1/n_{\text{large}})$*100 percentile, and that $n_{\text{small}} < n_{\text{large}}/2$, where n_{large} and n_{small} are the sizes of the large and the small populations, respectively. For populations with $n_{\text{large}} \geq 100$, the skill level must already be in the 99$^{\text{th}}$ percentile, and the percentile for the skill level increases rapidly to implausible levels as the population size increases.

For the case of Tasmania, where it is claimed that individuals sufficiently skilled to carry out the task of making clothing using bone points would be found in a population of 8,000 persons but would unlikely be found in a population of 4,000 persons, the skill level would have to at least be in the 99.9875$^{\text{th}}$ percentile of skill levels for this to occur, yet the task in question is that of making simple clothing using an easily made bone tool. The treadmill phenomenon occurs, then, if at all, only for small populations of tens of individuals at the most, and the only scenario where we could imagine that there would be small populations of individuals each with a specific skill would be sedentary communities with craft specialization in which a few artisans produce artifacts of a particular kind that are then distributed, by various means, to other group members.

The treadmill model implies that implements more complex than, for example, a bone point require even greater skills for their invention and for their implementation. These greater skills would only be found in yet far larger populations, according to the treadmill model, hence as we go from simple tools consisting of a single part to complex tools such as the harpoon with 33 parts made by the Inuit of eastern Greenland, the population size required to produce complex implements would quickly be vastly larger than the total number of hunter-gathers in a geographic region, as the complexity of implements increased. Not surprisingly, the empirical data on hunter-gatherer tool complexity has neither correlation with the census population size nor with the interaction population size. Thus, the prediction of the treadmill model that tool complexity is driven by increase in population density – which is a proxy measure for the interaction population size – is not supported empirically.

If anything, anthropological and ecological theorizing suggest that cultural complexity generates high population density rather than the other way around. There is likely such an effect playing out across time, but it is not a social effect per se and it has to do with population pressure (Keeley, 1988); i.e., to the ecological relation between the group and its environment and the competition between groups utilizing overlapping territories for resource procurement and/or resource production. Increase in complexity of implements through investment in technology points to an increase in the intensity with which local resources are procured and utilized. This may lead to replacement of neighboring low-complexity, low density and area-demanding groups. Alternatively, neighboring groups may develop more complex technology that allows them to tap into new and/or marginal resources (Read, 1987). The Broad-Spectrum Revolution (Flannery, 1969; Stiner, 2000; Stiner et al., 2000) exemplifies this process.

Our argument does not, however, deny that there are conditions under which increased technological complexity is dependent upon increased population size, just that this does not occur in the manner expressed in the treadmill model. The conditions relating increase in population size to technological complexity occur when the support system required for implementation of the more complex technological system requires a larger population, not necessarily because what is produced through the technology is complex, but because there are more subgoals that must first be implemented for the technology, itself, to be implemented and this requires more persons than are available in a smaller population. Situations like this arise with division of labor for implementation of the subgoals involve in technological production, ranging from resource procurement to systems of distribution, and are magnified with intensification of production. Conditions like this, however, do not occur in small hunter-gatherer societies.

References

Acerbi, A. & Mesoudi, A. (2015). If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution. *Biology & Philosophy* 30, 481-503.

Andersson, C. (2013). Fidelity and the emergence of stable and cumulative sociotechnical systems. PalaeoAnthropology 2013, 88–103.

Andersson, C. & Read, D. (2014). Group size and cultural complexity. Nature 511, http://dx.-doi.org/10.1038/nature13411.

Andersson, C. & Read, D. (2016). The evolution of cultural complexity, Not by the treadmill alone. Current Anthropology 57(3), 261-286.

Andersson, C., & Törnberg, P. (2018). Wickedness and the anatomy of complexity. Futures 95, 118–138.

Andersson, C., Törnberg, A. & Törnberg, P. (2014a). An evolutionary developmental approach to cultural evolution. *Current Anthropology* 55(2), 154–174. https://doi.org/10.1086/675692

Andersson, C., Törnberg, A., & Törnberg, P. (2014b). Societal systems: complex or worse? *Futures* 63, 145–157.

Andersson, C., Törnberg, A., & Törnberg, P. (forthcoming). Wicked Systems and the friction between “old theory and new data” in biology, social science and archaeology. In A. Love & W. C. Wimsatt (Eds.), *Beyond the meme: Development and structure in cultural evolution.* Minneapolis: University of Minnesota Press.

Asfaw, B., Beyene, Y., Suwa, G., Walter, R. C., White, T. D., Woldegabriel, G., & Yemane, T. (1992). The earliest Acheulean from Konso-Gardula. *Nature* 360, 732–734.

Avruch, K. (1998). *Culture and conflict resolution*. Washington DC: United States Institute of Peace Press.

Baldwin, M. J. (1896). A new factor in evolution. *The American Naturalist* 30(354), 441–451.

Barrett, L. & Henzi, P. (2005). The social nature of primate cognition. *Proceedings Biological Sciences / The Royal Society* 272(1575), 1865–1875.

Binford, L. R. (1980). Willow smoke and dog’s tails: Hunter-gatherer settlement systems and archaeological site formation. *American Antiquity* 45, 4–20.

Binford, L. R. (2001). *Constructing frames of reference.* Berkeley: University of California Press.

Boesch, C. & Boesch, H. (1983). Optimisation of nut-cracking with natural hammers by wild chimpanzees. *Behaviour* 83, 265-286.

Boyd, R. (2018). A different kind of animal: How culture transformed our species. Princeton: Princeton University Press.

Boyd, R. & Richerson, P. J. (1985). *Culture and the evolutionary process.* Chicago: University of Chicago Press

Bril, B., Roux, V., & Dietrich, G. (2005). Stone knapping: Khambhat (India) a unique opportunity? In V. Roux & B. Bril (Eds.), *Stone knapping:The necessary conditions for a uniquely hominin behaviour* (pp. 53–72). Cambridge: McDonald Institute for Archaeological Research.

Bubandt,N, & Otto, T. (2010). Anthropology and the predicaments of holism. In T. Otto & N. Bubandt (Eds.), *Experiments in holism* (pp. 1-15). London: Wiley-Blackwell.

Buchanan, B., O’Brien, M. J. & Collard, M. (2015). Drivers of technological richness in prehistoric Texas: An archaeological test of the population size and environmental risk hypotheses. *Archaeological and Anthropological Science.* DOI 10.1007/s12520-015-0245-4.

Burdett, E. R., Dean, R. L. G. & Ronfard, S. (2017). A diverse and flexible teaching toolkit facilitates the human capacity for cumulative culture. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-017-0345-4

Caldwell, C. A. & Millen, A. E. (2010). Human cumulative culture in the laboratory: effects of (micro) population size. *Learning and Behavior* 38(3), 310–318.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review* 67(6), 380–400.

Campbell, D. T. (1965). Variation and selective retention in socio-cultural evolution. In H. R. Barringer, G. I. Blanksten & R. W. Mack (Eds.), *Social change in developing areas: A reinterpretation of evolutionary theory* (pp. 19–48). Cambridge: Schenkman Publishing Company.

Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp. 413–463). La Salle: Open Court.

Castro, L. & Toro, M. A. (2014). Cumulative cultural evolution: The role of teaching. *Journal of Theoretical Biology* 347(1), 74–83.

Cavalli-Sforza, L. L. & Feldman, M. W. (1981). *Cultural transmission and evolution*. Stanford: Stanford University Press.

Chatters, J.C. (1995). Population growth, climatic cooling, and the development of collector strategies on the Southern Plateau, Western North America. *Journal of World Prehistory* (9)3, 341-400.

Collard, M., Buchanan, B., Morin, J. & Costopoulos, A. (2011a). What drives the evolution of hunter-gatherer subsistence technology? A reanalysis of the risk hypothesis with data from the Pacific Northwest. *Philosophical Transactions of the Royal Society B* 366, 1129–1138.

Collard, M., Buchanan, B., & O'Brien, M. J. (2013). Population size as an explanation for patterns in the Paleolithic archaeological record: More caution is needed. *Current Anthropology* 54, S388–S396.

Collard, M., Buchanan, B., O'Brien, M. J. & Scholnick, J. (2013a). Risk mobility or population size? Drivers of technological richness among contact-period western North American hunter-gatherers. *Philosophical Transactions of the Royal Society B* doi:10.1098/rstb.2012.0412

Collard, M., Buchanan, B., Ruttle, A. & O'Brien, M. J. (2011b). Niche construction and the toolkits of hunter-gatherers and food producers. *Biological Theory* 6(3), 251-259.

Collard, M., Kemery, M. D. & Banks, S. (2005). Causes of toolkit variation among hunter-gatherers: A test of four competing hypotheses. *Canadian Journal of Archaeology* 29, 1–19.

Collard, M., Ruttle, A., Buchanan, B. & O'Brien, M. J. (2013b). Population size and cultural evolution in non-industrial food-producing societies. *PLoS ONE* 8, e72628.

Collard, M., Vaesen, K., Cosgrove, R. & Roebroeks, W. (2016). The empirical case against the 'demographic turn' in Palaeolithic archaeology. *Philosophical Transactions of the Royal Society B* 371(1698), 20150242

Corbey, R. Jagich, A., Vaesen, K. & Collard, M. (2016). The Acheulean hand axe: More like a bird's song than a Beatles' tune? *Evolutionary Anthropology* 25, 6-19.

Costa-Peirce, B. A. (1985). Aquaculture in ancient Hawaii. *BioScience* 37, 320-331.

Csibra, G. & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences* 13(4), 148–153. http://doi.org/10.1016/j.tics.2009.01.005

Csibra, G. & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. Philosophical Transactions of the Royal Society of London Series B Biological Sciences 366(1567), 1149–1157.

Dean, L. G., Vale, G. L., Laland, K. N., Flynn, E. & Kendal, R. Dean, 2014L. (2014). Human cumulative culture: A comparative perspective. *Biological Reviews of the Cambridge Philosophical Society* 89 (2), 284–301.

D'Errico, F. & Banks, W. E. (2013). Identifying mechanisms behind Middle Paleolithic and Middle Stone Age cultural trajectories. *Current Anthropology* 54(S8), S371–S387.

Derex, M., Beugin, M-. P., Godelle, B. & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature* 503(7476), 389–391.

Derex, M., Beugin, M-. P., Godelle, B. & Raymond, M. (2014). Derex et al. reply. *Nature* 511, http://dx.doi.org/10.1038/nature13412.
Diamond, J. (1978). The Tasmanians: The longest isolation, the simplest technology. *Nature* 273, 185-186.
Diez-Martín, F., Sánchez, P., Yustos, D., Uribelarrea, E., Baquedano, D. F., Mark, A., Mabulla … & Domínguez-Rodrigo, M. (2015). The origin of the Acheulean: The 1.7 million-year-old site of FLK West Olduvai Gorge (Tanzania). *Scientific Reports* 5, 1–9.
Domínguez-Rodrigo, M., Serrallonga, J., Juan-Tresserras, J., Alcala, L. & Luque, L. (2001). Woodworking activities by early humans: A plant residue analysis on Acheulian stone tools from Peninj (Tanzania). *Journal of Human Evolution* 40(4), 289–299. http://doi.org/10.1006/jhev.2000.0466
Durham, W.H. (1991). Coevolution: Genes, culture, and human diversity. Stanford: Stanford University Press.
Eerkens, J. W. & Lipo, C. P. (2005). Cultural transmission copying errors and the generation of variation in material culture and the archeological record. *Journal of Anthropological Archaeology* 24, 316–334.
Eerkens, J. W. & Lipo, C. P. (2007). Cultural transmission theory and the archeological record: Providing context to understanding variation and temporal changes in material culture. *Journal of Archaeological Research* 15, 239–274.
El Guindi, F. (2015). Westermarck! Presentation at 114th Annual Meeting of the American Anthropological Association, Colorado Convention Center. Los Angeles: El Nil Research.
El Guindi, F. & Read, D. (2012). Westermarck hypothesis reconsidered: A comment on Kushnick and Fessler. *Current Anthropology* 53(1), 134–135.
Érdi, P. (2008). *Complexity explained*. Berlin Heidelberg: Springer Verlag.
Ferraro, G. (1998). *The cultural dimension of international business*. 3rd Edition. Upper Saddle River: Prentice Hall.
Firth, R. (1951). Elements of social organization. London: Watts.
Flannery, K. V. (1969). Origins and ecological effects of early domestication in Iran and the Near East In P. J. Ucko & G. W. Dimbleby (Eds.), *The domestication and exploitation of plants and animals* (pp. 73–100). Chicago: Aldine.
Foley, R. A. (1987). Hominid species and stone tools assemblages: How are they related? *Antiquity* 61, 380-392.
Fortes, M. (1969). Kinship and the social order: The legacy of Lewis Henry Morgan. Chicago: Aldine.
Galef, B. G. (1992). The question of animal culture. *Human Nature* 3(21):157-178.
Galef, B. G. & Laland, L. N. (2005). Social learning in animals: Empirical studies and theoretical models. BioScience 55(6), 489-499
García-Medrano, P., Ollé, A., Mosquera, M., Cáceres, I., Díez, C. & Carbonell, E. (2014). The earliest Acheulean technology at Atapuerca (Burgos Spain): Oldest levels of the Galería site (GII Unit). *Quaternary International* 353, 170–194. http://doi.org/10.1016/j.quaint.2014.03.053
Gärdenfors, P. & Högberg , A. (2017). The archaeology of teaching and the evolution of *Homo docens*. *Current Anthropology* 58(2), 188-208.

Gilligan, I. (2007). Another Tasmanian paradox: Clothing and thermal adaptations in Aboriginal Australia. BAR International Series 1710. Oxford: Archaeopress.

Gilligan, I. (2014). Clothing and modern human behavior: The challenge from Tasmania. In R. Dennell & M. Porr (Eds.), *Southern Asia, Australia and the search for human origins* (pp. 189-199). Cambridge: Cambridge University Press.

Goodenough, W. (Editor). (1964). Explorations in cultural anthropology: Essays in honor of George Peter Murdock. New York: McGraw-Hill Book Company.

Graves, P. (1994). Flakes and ladders: What the archaeological record cannot tell us about origins of language. *World Archaeology* 26, 158–171.

Greenberg, C. (1961). *Art and culture*. Boston: Beacon Press.

Hage, P. (1997). Unthinkable categories and the fundamental laws of kinship. *American Ethnologist* 24(3), 652-667.

Helgason, A., Pálsson, G., Pedersen, H. S., Angulalik, E. & Gunnarsdóttir, E. D. et al. (2006). mtDNA variation in Inuit populations of Greenland and Canada: Migration history and population structure. *American Journal of Physical Anthropology* 130, 123-134.

Helgason, A. (2010). Email communication to Dwight Read January 22 2010.

Henrich, J. (2004). Demography and cultural evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity* 69, 197-214.

Henrich, J. (2006). Understanding cultural evolutionary models: A reply to Read's critique. *American Antiquity* 71, 771-782.

Henrich, J., Boyd, R., Derex, M., Kline, K., Mesoudi, A., Muthukrishna, M., Powell, A., Shennan, S., & Thomas, M. (2016). Appendix to understanding cumulative cultural evolution: A reply to Vaesen, Collard, et al. Available at SSRN: https://ssrn.com/abstract=2798257 or http://dx.doi.org/10.2139/ssrn.2798257.

Henrich, J. & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* 22(3), 165-196.

Herbert, C. (1991). Culture and anomie: Ethnographic imagination in the Nineteenth Century. Chicago: University of Chicago Press.

Herrygers, C. (2002). A comparative analysis of wood residues on experimental stone tools and early stone age artifacts: A Koobi Fora case study a comparative analysis of wood residues on experimental stone tools. *McNair Scholars Journal* 6(1), 65–75.

Hodder, I. (2012). Entangled: An archaeology of the relationships between humans and things. New York: Wiley-Blackwell.

Hoijer, H. (1948). Linguistic and cultural change. *Language* 24(4), 335-345.

Jones, R. (1977) The Tasmanian paradox. In V.S. Wright (Ed.), *Stone tools as cultural markers: Change, evolution and complexity* (pp. 189-204). New Jersey: Humanities Press.

Keeley, L. H. (1988). Hunter-gatherer economic complexity and “population pressure”: A cross-cultural analysis. *Journal of Anthropological Archaeology* 7, 373–411.

Keeley, L. H. & Toth, N. (1981). Microwear polishes on early stone tools from Koobi Fora Kenya. *Nature* 293, 464–465.

Kempe, M. & Mesoudi, A. (2014). An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior* 35(4), 285–290.

Kline, M. A. (2010). Email communication to Dwight Read December 7, 2010.
Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences* 38, 1–70.
Kline, M. A. & Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B* 277, 2559-2564.
Kroeber, A.L. (1919). On the principle of order in civilization as exemplified by changes of fashion. *American Anthropologist* 21(3), 235-263.
Laland, K. N. & Janik, V. M. (2006). The animal cultures debate. *Trends in Ecology & Evolution* 21(10), 542-547.
Lane, D. A. (2016). Innovation cascades: artefacts organization and attributions. Philosophical Transactions of the Royal Society of London Series B Biological Sciences 371(1690), 20150194- https://doi.org/10.1098/rstb.2015.0194.
Lane, D., Maxfield, R. M., Read, D. & van der Leeuw, S. (2009). From population to organization thinking. In D. Lane, D. Pumain, S. E. van der Leeuw & G. West (Eds.), *Complexity perspective in innovation and social change* (pp. 11-42). Springer: Berlin.
Laubichler, M. D. & Maienschein, J. (2007). From embryology to evo-devo: A history of developmental evolution. (M. D. Laubichler & J. Maienschein Eds.). Cambridge: MIT Press.
Laubichler, M. D. & Maienschein, J. (2013). Developmental evolution. In M. Ruse (Ed.), *The Cambridge Encyclopedia of Darwin and evolutionary thought* (pp. 375–382). Cambridge: Cambridge University Press.
Laubichler, M. D. & Renn, J. (2015). Extended evolution: A conceptual framework for integrating regulatory networks and niche construction. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 324B(7), 565–577.
Leavitt, G. C. (2005). Incest and inbreeding avoidance: A critique of Darwinian social science. Lewiston: Edwin Mellen Press.
Leaf, M. & Read, D. (2012). Human thought and social organization: Anthropology on a new plane. Lanham: Lexington Press.
Lycett, S. J. & von Cramon-Taubadel, N. (2015). Toward a "quantitative genetic" approach to lithic variation. *Journal of Archaeological Method and Theory* 22(2), 646–675.
Marcot, J. D. & McShea, D. W. (2007). Increasing hierarchical complexity throughout the history of life, phylogenetic tests of trend mechanisms. *Paleobiology* 33(2), 182–200.
Matsuzawa, T. (1996). Chimpanzee intelligence in nature and in captivity: Isomorphism of symbol use and tool use. In W. C. McGrew, L. F. Marchant & T. Nishida (Eds.), *Great ape societies* (pp. 196-209). Cambridge: Cambridge University Press.
Matsuzawa, T. (2007). Comparative cognitive development. *Developmental Science* 10, 97-103.
McGrew, W. C., Ham, R. M., White, L. J. T., Tutin, E. G. & Fernandez, M. (1997). Why don't chimpanzees in Gabon crack nuts? *International Journal of Primatology* 18, 353-374.
McShea, D. W. (1991). Complexity and evolution: what everybody knows. *Biology and Philosophy* 6, 303–324.
McShea, D. W. (2000). Functional complexity in organisms: parts as proxies. *Biology and Philosophy* 15, 641–668.
Middleton, R. (1962). Brother-sister and father-daughter marriage in ancient Egypt. *American Sociological Review, 27,* 603-611.

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63(2), 81–97.

Mundt, T. (1857). *Pariser Kaiser-Skizzen*. 2 vols. Berlin: Verlag von Otto Janke.

Muthukrishna, M., Shulman, B. W., Vasilescu, V. & Henrich, J. (2014). Sociality influences cultural complexity. *Proceeding of the Royal Society of London B: Biological Sciences* 281(1774), 20132511.

Mulvaney, J. & Kamminga, J. (1999). *Prehistory of Australia*. Washington: Smithsonian Institution.

O'Brien, M. J. & Bentley, R. A. (2011). Stimulated variation and cascades, Two processes in the evolution of complex technological systems. *Journal of Archaeological Method and Theory* 18(4), 309–335.

O'Brien, M. J., Lyman, R. L., Mesoudi, A. & VanPool, T. L. (2010). Cultural traits as units of analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1559), 3797–3806.

Oswalt, W. H. (1976). An anthropological analysis of food-getting technology. New York: Wiley.

Petersen, R. (1984). East Greenland before 1950. In D. Damas (Ed.), *Handbook of North American Indians. Volume 5* (pp. 622-639). Washington DC: Smithsonian Institution

Petraglia, M. D. & Shipton, C. (2008). Large cutting tool variation west and east of the Movius Line. *Journal of Human Evolution* 55(6), 962–966.

Ping, K. (1999). Cultural presuppositions and misreadings. *Meta: Théorie et pratique de la traduction en Chine.* 44(1), 133–143.

Powell, A., Shennan, S. J. & Thomas, M. G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science* 324, 1298–1301. doi:10.1126/science.1170165

Querbes, A., Vaesen, K.& Houkes, W. (2014). Complexity and demographic explanations of cumulative culture. *PloS ONE* 9(7),e102543.

Read, D. (1984) An algebraic account of the American kinship terminology. *Current Anthropology* 25, 417-440.

Read, D. (1987). Foraging society organization, A simple model of a complex transition. *European Journal of Operational Research* 30, 320-326.

Read, D. (2001). What is Kinship? In R. Feinberg & M. Ottenheimer (Eds.), The cultural analysis of kinship: The legacy of David Schneider and its implications for anthropological relativism (pp. 78-117). Urbana: University of Illinois Press.

Read, D. (2006). Tasmanian knowledge and skill: Maladaptive imitation or adequate technology. *American Antiquity* 73, 599–625.

Read, D. (2007a). Artefact classification: A conceptual and methodological approach. Walnut Creek: Left Coast Press.

Read, D. (2007b). Kinship theory: A paradigm shift. *Ethnology* 46(4), 329-364

Read, D. (2008). An interaction model for resource implement complexity based on risk and number of annual moves. *American Antiquity* 73, 599 – 625.

Read, D. (2011). The misuse of a mathematical model: The Tasmanian case (reply to Henrich's response). *UCLA: Human Complex Systems.* Retrieved from https://escholarship.org/uc/item/88k8g4rj

Read, D. (2012a). How culture makes us human: Primate Social Evolution and the *Formation of Human Societies*. Walnut Creek: Left Coast Press.

Read, D. (2012b). Population size does not predict artifact complexity: Analysis of data from Tasmania Arctic hunter-gatherers and Oceania fishing groups. *UCLA: Human Complex Systems*. Retrieved from https://escholarship.org/uc/item/61n4303q.

Read, D. (2014a). Incest taboos and kinship: A biological or a cultural story? *Reviews in Anthropology* 43(12), 150-175.

Read, D. (2014b). The substance of cultural evolution: Culturally framed systems of social organization: Comment on "Smaldino 2014 The cultural evolution of emergent group-level traits." *Behavioral and Brain Sciences* 37(3), 270-271.

Read, D. (2017). "From past to present: kinship and the mission of anthropology," Plenary Talk XII Congress of Anthropologists and Ethnographers of Russia Izhevsk Russia July 5, 2017.

Read, D. (2018). "The deep history of kinship takes us back to procreation, but procreation is not the story of kinship." Paper given at the *Atelier d'Analyse Anonyme*, March 30-31, 2018, Paris. in the session titled: Peut-on penser á la parenté sans songer á la procreation?

Read, D., Fischer, M. D. Chit Hlaing, F. K. L. (2014). The cultural grounding of kinship: A paradigm shift. *L'Homme* 210(2), 63-89.

Read, D., Lane, D. & van der Leeuw, S. (2009). The innovation innovation. In D. Lane, D. Pumain, S. E. van der Leeuw & G. West (Eds.), *Complexity perspective in innovation and social change* (pp. 43-84). Berlin: Springer.

Read, D. & van der Leeuw, S. (2008). Biology is only part of the story … . *Philosophical Transactions of the Royal Society B* 363, 1959-1968.

Richerson, P. J. & Boyd, R. (2005). Not by genes alone: How culture transformed human evolution. Chicago: University of Chicago Press.

Rouse, I. (1939). *Prehistory in Haiti: A study in method*. Yale University Publications in Anthropology Number 21. New Haven: Yale University Press.

Ruff, C. B., Trinkhaus, E. & Holliday, T. W. (1997). Body mass and encephalization in Pleistocene Homo. *Nature* 387, 173–176. doi:10.1038/387173a0

Sanz, C. M. & Morgan, D. B. (2007). Chimpanzee tool technology in the Goualougo Triangle Republic of Congo. *Journal of Human Evolution* 52, 420–433.

Scheidel, W. (2005). Ancient Egyptian sibling marriage and the Westermarck effect. In A. P. Wolf , & W. H. Durham (Eds.), *Inbreeding, incest, and the incest taboo* (pp. 93-108). Palo Alto: Stanford University Press.

Shennan, S. J. (2001). Demography and cultural innovation: A model and some implications for the emergence of modern human culture. *Cambridge Archeological Journal* 11, 5–16. doi: 10.1017/s0959774301000014

Shepher, J. (1983). *Incest. A biosocial view*. New York: Academic Press.

Shils, E. (1981). *Tradition*. Chicago: University of Chicago Press.

Shor, E., & Simchai, D. (2009). Incest avoidance, the incest taboo, and social cohesion: Revisiting Westermarck and the case of the Israeli kibbutzim. *American Journal of Sociology* 117, 1803–1842.

Shor, E., & Simchai, D. (2012). Exposing the myth of sexual aversion in the Israeli kibbutzim: A challenge to the Westermarck hypothesis. *American Journal of Sociology* 117(5), 1509–1513.

Shott, M. (1986). Technological organization and settlement mobility: An ethnographic examination. *Journal of Anthropological. Research* 42, 15–51.
Shields, W. M. (1982). Philopatry, inbreeding, and the evolution of sex. Albany: State University of New York Press.
Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society* 6(106), 467–482.
Smaldino, P. (2014). The cultural evolution of emergent group-level traits. *Behavioral and Brain Sciences* 37(3), 243-295.
Sterelny, K. (2004). A review of “Evolution and Learning: The Baldwin Effect Reconsidered, edited by B. Weber and D. Depew.” *Evolution and Development* 6(4), 295–300.
Sterelny, K. (2011). Evolvability reconsidered. In B. Calcott & K. Sterelny (Eds.), *The major transitions in evolution revisited* (pp. 83–100). Cambridge: MIT Press.
Stiner, M. C. (2001). Thirty years on the “broad spectrum revolution” and Paleolithic demography. *Proceedings of the National Academy of Sciences of the United States of America* 98(13), 6993–6996.
Stiner, M. C., Munro, N. D. & Surovell, T. A. (2000). The tortoise and the hare. *Current Anthropology* 41(1), 39–73.
Sutton, M.Q. (2000). Strategy and tactic in the analysis of archaeological hunter-gatherer systems. *North American Archaeologist* 21(3), 217-231.
Taylor, T. (2010). The artificial ape: How technology changed the course of human evolution. New York: Palgrave Macmillan.
Taylor, W. (1948). *A study of archeology*. Memoir 69, American Anthropological Association. Reprinted, Carbondale: Southern Illinois University Press, 1967.
Tehrani, J. J. & Riede, F. (2008). Towards an archaeology of pedagogy: Learning teaching and the generation of material culture traditions. *World Archaeology* 40(3), 316–331.
Tomasello, M. (1996). Do apes ape? In C. M. Heyes & B. G. Galef Jr. (Eds.), *Social learning in animals: The roots of culture* (pp. 319-346). San Diego: Academic Press.
Torrence, R. (1983). Time budgeting and hunter-gatherer technology. In P. Halstead & J. O’Shea (Eds.), *Bad year economics: Cultural responses to risk and uncertainty* (pp. 11–22). Cambridge: Cambridge University Press.
Torrence, R. (1989). Re-tooling: towards a behavioral theory of stone tools. In R. Torrence (Ed.), *Time energy and stone tools* (pp. 57–66). Cambridge: Cambridge University Press.
Torrence, R. (2000). Hunter–gatherer technology: macro- and microscale approaches. In C. Panter-Brick, R. H. Layton, & P. Rowley-Conwy (Eds.), *Hunter–gatherers: an interdisciplinary perspective* (pp. 99–143). Cambridge: Cambridge University Press.
Turner, J. & Maryanski, A. (2009). *Incest: Origins of the taboo*. Boulder: Paradigm.
Tylor, E. B. (1929[1871]). *Primitive culture*. London: J Murray.
Vaesen, K. Collard, M., Cosgrove, R. & Roebroeks, W. (2016). Population size does not explain past changes in cultural complexity. *Proceedings of the National Academy of Sciences USA* 113(16), E2241-E2247.
Watson, P. (1995). Archaeology, anthropology, and the culture concept. *American Anthropologist* 97(4), 683-694.
Westermarck, E. (1922). The history of human marriage: Vol. 2. New York: Macmillan.

Wimsatt, W. C. (1975). Complexity and organization. In *Topics in the Philosophy of Biology* (pp. 174–193). Dordrecht Holland: D Reidel Publishing Company.

Wimsatt, W. C. (1999). Genes memes and cultural heredity. *Biology and Philosophy* 14, 279–310.

Wimsatt, W. C. (2013). Entrenchment and scaffolding: An architecture for a theory of cultural change. In L. R. Caporael, J. R. Griesemer & W. C. Wimsatt (Eds.), *Developing scaffolding in evolution, cognition and culture* (pp. 77–105). Cambridge: MIT Press.

Wimsatt, W. C. (2015). Entrenchment as a theoretical tool in evolutionary developmental biology. In A. Love (Ed.), *Conceptual change in biology: Boston studies in the philosophy and history of science* (pp. 365–402). Springer Netherlands. http://doi.org/10.1007/978-94-017-9412-1_17.

Wimsatt, W. C. & Griesemer, J. R. (2007). Reproducing entrenchments to scaffold culture: The central role of development in cultural evolution. In R. Sansom & R. N. Brandon (Eds.), *Integrating evolution and development: From theory to practice* (pp. 227–323). Cambridge: MIT Press.

Wolf, A. (1995). Sexual attraction and childhood association. A Chinese brief for Edward Westermarck. Stanford: Stanford University Press.

Wolf, A. (2014a). *Incest avoidance and the incest taboos*. Stanford: Stanford University Press.

Wolf, A. (2014b). Personal communication. Email from Arthur Wolf to Dwight Read dated 6/25/2014.

Wynn, T. (2002). Archaeology and cognitive evolution. *Behavior and Brain Science* 25, 389–438. doi:10.1017/S0140525X02530120.

Biographies

Dwight W. Read is a Distinguished Professor of Anthropology at UCLA, Emeritus, and a Fellow of the American Association for the Advancement of Science. He received his Ph.D at UCLA in mathematics, with research focus on abstract algebras. He does research and publishes in all of the subdisciplines of Anthropology. His current research considers the interrelationship between the material and the ideational domains in human societies. He has recently published three books: *Artifact Classification: A Conceptual and Methodological Approach* (2007, Left Coast Press), *How Culture Makes Us Human* (Series: Big Ideas in Little Books, 2012, Left Coast Press) and *Human Thought and Social Organization: Anthropology on a New Plane* with Murray Leaf (2012, Lexington Press), as well as more than 200 journal articles and book chapters and more than 150 national and international talks.

Claes Andersson is a Senior Researcher and Associate Professor in Complex Systems at Chalmers University of Technology in Gothenburg, Sweden, and an external fellow of the European Centre for Living Technology in Venice, Italy. He works synthetically and interdisciplinary on large-scale and long-term patterns and mechanisms of societal change common to complexity science, anthropology, biology and geography. Most of his research has been on human evolution and pre-history, and on urban and regional dynamics in the present and recent history.